Proceedings of the
9th Annual South-East European
Doctoral Student Conference

Edited by

R. Dautov, P. Gkasis, A. Karamanos, T. Lagkas, A. Prodromidou,
A. Ypsilanti

Preface

These proceedings represent the work of contributions to the ninth (9th) Annual SEERC Doctoral Student Conference (DSC2014) hosted by the University of Sheffield International Faculty, CITY College and organised by the South-East European Research Centre - Thessaloniki, Greece.

DSC2014 has grown and continues to evolve. For the last 9 years, the key aim remains to provide an opportunity for PhD students and young researchers to receive advice from experts in their chosen field of research. Having identified academic isolation as a problem that many doctoral students face today, SEERC aims to bring researchers together for establishing collaborative links between disciplines, for testing the ground for innovative ideas and for engaging the wider academic community.

Building on the success of the past eight conferences, this year's conference attracted a large number of submissions resulting in 55 presentations of full papers. The audience of the conference expanded beyond the boundaries of South-East Europe confirming the need for Doctoral Students to come together, discuss their experiences and gain external feedback to their work as well as listen to the progress and methodology of fellow PhD candidates.

These papers represent research from Albania, Armenia, Australia, Bulgaria, Czech Republic, Estonia, Former Yugoslav Republic of Macedonia, Germany, Greece, Italy, Kosovo (under UNSC 1244), Montenegro, Poland, Romania, Serbia, Spain, Sri Lanka, Turkey, United Kingdom, United States.

I hope that you enjoy the conference.

September 2014
Prof. Panayiotis H. KETIKIDIS
Chairperson of DSC2014

Conference Organisation

DSC2014 is organized by the South-East European Research Centre (SEERC), an overseas research centre of the University of Sheffield, established as a non-profit legal entity in Thessaloniki, Greece. SEERC was founded by CITY College, the University's International Faculty, in 2003.

Organising Committee

Ms. Christina Batzaka, *South-East European Research Centre*
Mr. Rustem Dautov, *South-East European Research Centre*
Dr. Pavlos Gkasis, *The University of Sheffield International Faculty, CITY College*
Ms. Anka Gorgiev, *South-East European Research Centre*
Ms. Athanasia Halatzouka, *South-East European Research Centre*
Dr. Anastasios Karamanos, *The University of Sheffield International Faculty, CITY College*
Dr. Thomas Lagkas, *The University of Sheffield International Faculty, CITY College*
Dr. Alexandra Prodromidou, *The University of Sheffield International Faculty, CITY College*
Ms. Mariana Tsatali, *South-East European Research Centre*
Dr. Antonia Ypsilanti, *The University of Sheffield International Faculty, CITY College*

Scientific Committee Chairs

Enterprise, Innovation and Development

Dr. Pavlos Gkasis, *The University of Sheffield International Faculty, CITY College*
Dr. Anastasios Karamanos, *The University of Sheffield International Faculty, CITY College*
Dr. Nick Williams, *The University of Sheffield*

Information and Communication Technologies

Dr. Thomas Lagkas, *The University of Sheffield International Faculty, CITY College*

Society and Human Development: Psychology, Politics, Sociology, and Education

Dr. Alexandra Prodromidou, *The University of Sheffield International Faculty, CITY College*
Dr. Antonia Ypsilanti, *The University of Sheffield International Faculty, CITY College*

Scientific Committee

Enterprise, Innovation and Development

Dr. Dilay Çelebi, *Istanbul Technical University*
Dr. Alok Choudhary, *The University of Sheffield*
Dr. Ana Cruz, *The University of Sheffield International Faculty, CITY College*
Dr. Gurdal Ertek, *Sabanci University*
Dr. Thanassis Fassas, *The University of Sheffield International Faculty, CITY College*
Dr. Andrea Genovese, *The University of Sheffield*
Mr. Petros Golitsis, *The University of Sheffield International Faculty, CITY College*
Dr. Alexandros Kapoulas, *The University of Sheffield International Faculty, CITY College*
Dr. Anastasios Karamanos, *The University of Sheffield International Faculty, CITY College*
Prof. Paul Latreille, *The University of Sheffield*
Dr. Milos Milovančević, *The University of Niš*
Mr. Fotis Missopoulos, *The University of Sheffield International Faculty, CITY College*
Dr. Nikolaos Mylonidis, *University of Ioannina*
Dr. Dimitrios Nikolaidis, *The University of Sheffield International Faculty, CITY College*
Dr. Konstantinos Priporas, *Middlesex University London*
Dr. Alexandros Psychogios, *The University of Hull*
Dr. Leslie T. Szamosi, *The University of Sheffield International Faculty, CITY College*
Dr. Željko Tekić, *The University of Novi Sad*
Dr. Robert Wapshott, *The University of Sheffield*
Dr. Anna Watson, *University of Hertfordshire*
Dr. Nick Williams, *The University of Sheffield*
Dr. Alexandros Zangelidis, *University of Aberdeen*
Prof. Siniša Zarić, *The University of Belgrade*

Information and Communication Technologies

Dr. José María Álvarez Rodríguez, *Universidad Carlos III de Madrid*
Prof. Costin Bădică, *University of Craiova*
Prof. Tudor Balanescu, *University of Piteşti*
Prof. Ömer Faruk Bay, *Gazi University*
Assoc. Prof. Alexander Chatzigeorgiou, *University of Macedonia*
Prof. Betim Çiço, *South East European University*
Dr. Tony Cowling, *The University of Sheffield*
Dr. Konstantinos Dimopoulos, *The University of Sheffield International Faculty, CITY College*
Prof. Christos Douligeris, *University of Piraeus*
Dr. Dimitris Dranidis, *The University of Sheffield International Faculty, CITY College*
Dr. George Eleftherakis, *The University of Sheffield International Faculty, CITY College*
Dr. Anabel Fraga, *Universidad Carlos III de Madrid*
Prof. Neki Frasheri, *Polytechnic University of Tirana*
Assoc. Prof. Akli Fundo, *Polytechnic University of Tirana*
Assist. Prof. Mathias Funk, *Eindhoven University of Technology*
Dr. Christos Georgiadis, *University of Macedonia*
Dr. Marian Gheorghe, *The University of Sheffield*
Prof. Florin Gorunescu, *University of Craiova*
Dr. Fabrizio Granelli, *University of Trento*
Mr. Thanos Hatziapostolou, *The University of Sheffield International Faculty, CITY College*
Prof. Florentin Ipate, *University of Piteşti*
Mr. Isaak Kavasidis, *South-East European Research Centre*
Prof. Petros Kefalas, *The University of Sheffield International Faculty, CITY College*
Prof. Evangelos Kehris, *Technologal Education Institute of Serres*
Dr. Thomas Lagkas, *The University of Sheffield International Faculty, CITY College*
Prof. Juan Llorens, *Universidad Carlos III de Madrid*
Dr. Valentín Moreno, *Universidad Carlos III de Madrid*
Dr. Iraklis Paraskakis, *South-East European Research Centre*
Assist. Prof. Biljana Risteska Stojkoska, *Ss. Cyril and Methodius University*
Dr. Ilias Sakellariou, *University of Macedonia*
Prof. Demetrios Sampson, *University of Piraeus*
Dr. Anthony Simons, *The University of Sheffield*
Mrs. Anna Sotiriadou, *The University of Sheffield International Faculty, CITY College*
Dr. Ioanna Stamatopoulou, *The University of Sheffield International Faculty, CITY College*
Dr. Mike Stannett, *The University of Sheffield*
Dr. Anca Vasilescu, *Transilvania University of Brasov*

Dr. Hüsnü Yenigün, *Sabanci University*
Prof. Katerina Zdravkova, *Ss. Cyril and Methodius University*

Society and Human Development: Psychology, Politics, Sociology, and Education

Dr. Başak Alpan, *Middle East Technical University*
Dr. Panagiotis Bamidis, *Aristotle University of Thessaloniki*
Dr. Odeta Barbullushi, *European University of Tirana*
Prof. Stefan Bouzarovski, *University of Manchester*
Ms. Agnes Brunet, *The University of Sheffield International Faculty, CITY College*
Dr. Elissavet Chrysochoou, *The University of Sheffield International Faculty, CITY College*
Dr. Angeles F. Estevez, *Universidad de Almeria*
Dr. Tonin L. Gjuraj, *European University of Tirana*
Dr. Pavlos Gkasis, *The University of Sheffield International Faculty, CITY College*
Dr. Elena Griva, *University of Western Macedonia*
Dr. Aristea Ladas, *The University of Sheffield International Faculty, CITY College*
Dr. Lambros Lazuras, *The University of Sheffield International Faculty, CITY College*
Dr. Rod Nicholson, *The University of Sheffield*
Dr. Saska Petrova, *University of Manchester*
Dr. Suzie Savvidou, *The University of Sheffield International Faculty, CITY College*
Dr. Klio Semoglou, *University of Western Macedonia*
Dr. Nikolaos Tsigilis, *Aristotle University of Thessaloniki*
Dr. Ana Vivas, *The University of Sheffield International Faculty, CITY College*
Dr. Alexander Wochnik, *University of Manchester*

Sponsoring Institutions

South-East European Research Centre (SEERC), Greece
The University of Sheffield International Faculty, CITY College, Greece
The University of Sheffield, United Kingdom

Table of Contents

Enterprise, Innovation and Development

Information and Communication Technologies

Society and Human Development Psychology, Politics, Sociology and Education

Enterprise, Innovation and Development

Globalization: Managing Multinationals – A comparison of offshoring and outsourcing strategies in UK and German multinational corporations

Anthony Mitchell[1]

[1]Part Time PhD Student, University of Hertfordshire, Business School, Hatfield UK

`anthony@mitchellpalomares.com`

Abstract. This paper outlines research exploring differences in approach to outsourcing and offshoring by multinational corporations. In particular, a comparison is drawn between German and UK headquartered organizations in the airline and engineering sectors. A novel conceptual framework is developed; and the initial findings from two case studies are that German organizations are less inclined to outsource (in both sectors) preferring to retain control as a wholly owned business offshore. The UK businesses were less risk adverse and seemed to be more flexible and agile in their sourcing policies – using competitors when there is a sound business case. The relationships' with trade unions / works council was also found to be very different, with reluctance by management in Germany to progress radical initiatives. A favorable economy in Germany has also created an environment in which overseas expansion could take place without a loss of jobs at home.

Keywords: *offshore, outsource, varieties of capitalism*

1   Introduction

Offshoring and outsourcing represent on-going and accelerating (at least until recently) trend in the reorganization and restructuring of firms and has become a major part of (although not an exclusive driver of) the globalization trend. Offshoring can be defined as the performance of tasks in a different country to that where the firm's headquarters is located; while outsourcing may be regarded as the performance of tasks under some contractual arrangement by an unrelated third party (Harms, Lorz et al, 2009). Mergers and acquisition have a high risk of failure (Mitchell, 2004) and in recent years organizations have therefore sought alternative means of non-organic growth such as partnerships, joint ventures and alliances (Barthélemy, 2011). While the initial justification to offshore is typically to arbitrage labor costs, the rapid growth in demand for outsourcing may lead to cost increases (Economist, 2011) and justification increasingly becomes a complex balance of proximity to markets, suppliers, ability to innovate and institutional factors such as governance and immigration policy (Pisano, 2009). Further, there is an increasing trend to outsource and offshore activities that demand higher levels of skills. According to Kirkegaard (2008) few topics in

international economics have risen faster to the top of the political agenda, while also being so poorly understood and quantified as has outsourcing. Recent economic pressures have led governments in the United States and Europe to 'encourage' multinationals to return jobs and investment back to home markets (Economist, 2011); beyond this, backshoring and reverse offshoring have been motivated by poor or disappointing experiences in host countries, and declining economic conditions at home (Liebl, Morefield et al, 2010).

However, the institutional aspects of offshoring are under-explored and this research aims to compare the practices, strategies and outcomes for case study firms from The UK and Germany, which are characterized by different capitalist models (Hall and Soskice, 2001; Lane, 1998). It is suggested that German firms for example, typically have stronger institutional links than typical UK competitors (Lane 2006 cited in Morgan, Whitley and Moen, 2006). Furthermore, UK and German economies have different comparative advantages and industrial infrastructures, yet both countries also play host to a number of successful multinationals (MNC). The institutional context here can be understood as *both* the configuration of formal institutions (government, banks, trade unions and other firms) or as deeply embedded business practices and norms and 'ways of doing business'. This will shed light on how UK and German competing organizations differ in managing global expansion, and take advantage of the various resources and support available.

Following German reunification (1990) a period of austerity and strict wage control took place in Germany, and this helped to drive investment at home together with a strong export led economic revival. In 2012 German productivity was assessed to be 24 percentage points ahead of the UK in terms of output per hour (O'Connor, 2013). UK productivity is also currently 16 percentage points below the G7 average – the widest gap since 1994. A contested area is that the UK has been retaining employees rather than losing jobs to offshoring, while new work is created by UK outsourcing providers (see below). Throughout the 2008-9 recession, increased part-time working in the UK and even the hiring of new employees occurred at a time of minimal growth (O'Connor, 2013).

This research should be of interest to researchers, students and business managers. Also to those who are interested in globalization, the role of the multinational corporation, the relationship between a headquarters and its divisional or national subsidiaries. A further challenge is the extent to which offshoring and outsourcing practices have created wealth for shareholders, the host country and employees. Added interest is generated by challenging popular questions and criticism made of multinationals and their role in globalization together with the debate by politicians and others on policy towards domestic employment and wealth creation at home at a time of prolonged economic uncertainty.

1.1   The overall aim of the research is:

To examine the extent to which the offshoring and outsourcing strategies of UK and German based multinational corporations (MNCs) are embedded in the institutional

contexts of their respective home countries. This gives rise to a number of sub – questions:

1. What are the differences between UK and German based MNCs in the geographical, functional and temporal patterns of outsourcing and offshoring?

2. How far do mechanisms such as ownership, control, coordination and the degree of autonomy differ between the UK and Germany?

3. How is this reflected in divergent international divisions of labour regarding the employment of indigenous or ex-pat managers from the home country?

4. To what extent do preferences for cultural proximity affect location choices?

5. What is the influence of trade unions in the process of outsourcing and offshoring and how is this reflected in the structuring of the firms' labour markets?

6. What evidence is there of a reversal in policy – backshoring / reversed offshoring / outsourcing and why may it be occurring?

2   Literature Review

The purpose of the first contextual stage of the literature review is to review the varying definitions, challenges with measurement, recent trends, background issues to, and the debate around outsourcing and offshoring. This will help in understanding the motivation for offshoring and outsourcing. Firstly, some definitions because the two terms outsourcing and offshoring are sometimes confused and deployed in very different scenarios. This will provide a context for the changes that have been taking place at the level of a firm in response to globalization and competition.

Offshoring means that work is moved outside the home country and therefore has geographical connotations, usually to a country which can perform the work at lower cost, or perhaps has special skills; although there might also be a business case for offshoring around new market entry and moving operations closer to the country of destination.

Outsourcing currently implies that an organization decides to move selected activities from in-house (inside the organization) to a third party or external supplier through a formal contract arrangement. The supplier may or may not be in the same country of origin as the organization undertaking the outsourcing. The reasons for doing this may be multiple, but the usual starting point is to reduce costs, often labour and associated overhead charges. In so doing, the instigating organization can be said to be reorganising its value chain and moving either core or support activities to the responsibility of another organization.

 Measurement difficulties often arise from problems associated with the identification beforehand and the allocation of costs and/or poor recording of government statis-

tics. Offshoring work in particular may be outsourced to a third party or indeed under-taken through a wholly owned subsidiary business (adapted from Contractor, 2010). Questions continue to be raised about the value of multinational expansion (Contractor, 2012). While sourcing costs may be reduced locally, and foreign knowledge and intellectual property may be acquired in rapidly developing markets as can the hedging of currency risks. There are a number of other costs to consider; e.g. R&D and headquarter costs, often retained in the home market may increase substantially. Each foreign affiliate may have to incur substantial reorganization costs and change for example to incorporate group information and accounting systems, there may also be increased overheads to facilitate group controls and quality systems. Central costs of coordination will increase as the number of foreign markets rise, along with supply chain and inventory costs, risks of stock-out, supply failures. Institutional and cultural distance issues again add complexity, communication challenges and potential cost.

Offshoring and outsourcing could be analyzed as global disaggregation of the value chain and as an attempt to combine comparative advantages of geographic location with an organization's resources and competencies to maximise competitive advantage (Mudambi, 2010). The interplay of comparative and competitive advantages determines the optimal location of value chain components (offshoring decisions) as well as the boundaries of the firm and the control strategy (outsourcing decisions).

Three different but interrelated strands of theory have also been explored. From the fields of:

1. Business and economics, (Barney,1991) the Resource Based View (RBV).
2. Geography and economics, (Hall and Soskice, 2001) the concept of differing Varieties of Capitalism (VoC); and finally, from
3. Operations, geography, economics and strategy, (Coe et al, 2004) the concept of Global Production Networks (GPN).

The intention is to synthesize these differing approaches together with an understanding of offshoring to answer the research questions and to explore differences in how German and UK multinationals operate in specific business sectors, and manage offshoring / outsourcing processes in particular. This will also help in developing a conceptual framework – explored further under Methodology.

The lack of research on the interdependencies of geography and control is under-played considering that firms operating in international markets face these decisions simultaneously (Dunning, 1988) and so whilst addressed in part by researchers of GPNs, the field is contested. Making these decisions independent of each other leads to short term, tactical sub-goal optimization. The strategic integration of these decisions can result in significant firm-level performance improvements (Banker et al., 1984). Most of the offshoring literature takes control decisions as a given. Similarly, the mainstream literature on outsourcing usually fails to explore the location decision.

Understanding the cost-benefit of offshoring and outsourcing is informed by RBV theory and concepts. This goes beyond the simple assumption of labor cost arbitrage towards the complexities of disaggregating home based processes and deciding what exactly to move offshore and where to locate it. Behavior, whether rational or not, can be explored between buyers, suppliers and third parties in negotiating contracts and rents. If this can be combined with a better understanding of how to ensure that eco-

nomic goals are embedded into social structures and the subsequent impact on behaviour then we have a compelling approach.

There are obvious limitations in clustering nation states, nevertheless broad comparisons seem possible. VoC can provide fascinating insights to the role of governments and institutions in juggling support and resources from the public to the private sector (and vice versa) also the extent to which institutions or the market influence prices and positioning. The real issue is the extent to which this benefits longer term growth and prosperity for firms and their shareholders. Whether coordinated versus liberal, production versus finance dominated, or corporatist versus pluralist private enterprise, most writers on VoC agree on distinct differences between UK and German systems of capitalism. The significant distinction is how German or UK MNCs then coordinate policy and whether they take their lead from the market or influential institutions to coordinate stakeholders. Further understanding of inter-firm linkages, power and competition is provided by the study of GPNs. The role of the lead firm is considered crucial in managing the impact of institutional policy on resource allocation decisions. Once offshore processes are sufficiently embedded that they add value back to the lead firm, further complex decisions are often required on (re)positioning (typically expensive) R&D and innovation resources, along with suppliers and customer markets. There seem to be several issues that are underplayed by existing literature.

Firstly, institutional aspects of differing workplace environments and management groups largely responsible for decision making and policy setting of outsourcing and offshoring activity. If we consider the lead firm in a GPN, then there is an attractive argument that sustainable competitive advantage depends upon the firm's ability to manage the institutional context of its resource decisions (Oliver, 1997). Hence combining the resource based view with institutional perspectives from organizational theory overcomes both some of the criticism of VoC (Granovetter, 1992) and seems compelling in practice. Institutional theory assumes that individuals are motivated to respond to external pressures. A criticism of GPN research (Hess and Wai-chung Yeung, 2006) is that empirical studies have a preference for qualitative interviews with actors rather than empirical research data on the mechanisms and processes of GPNs. The 'cultural clash' that arose from European post socialist transformation over the past 17 years has attracted the attention of business partners from across the CEE. The body of organizational knowledge based on traditional, stable western market economies needs rethinking for sometimes unstable and ambiguous post- socialist environments (Soulsby and Clark, 2007). State Owned Enterprises (SOE's) tend to have functional hierarchies designed to have instructions and targets handed down through the various levels. A well connected MNC and the use of FDI could be critical in changing past practice and delivering demanding service level agreements (SLA's). This is clearly not easy, local managers will criticize the economic rationality of western values and practices such as financial control and downsizing while trying to defend local values such as a duty of care and the value of labour (Soulsby and Clark, 2007).

Secondly, a hotly contested area includes groups of labor and the impact of offshoring on employment levels. It has been suggested that improvements in technology (that link tasks across distance and borders) lead to domestic job losses through off-

shoring but also create jobs from cost savings associated with enhanced trade. Employment takes time to adjust to improvements in offshoring technology (Kohler and Wrona, 2010). In support, assume that an organization relocates 500 jobs to India than this constitutes a relocation effect (Gorg, 2011). If however, offshoring these jobs results in an increase in business productivity and sales increase so does employment. So whilst there may well be contested arguments for and against offshoring with disputes on the pros and cons there is also a level of misreporting which confuses the facts. This is interesting to note as data reported tends to focus on jobs lost through offshoring misrepresenting the true effect; reconciling jobs lost and new jobs created (elsewhere) is extremely difficult. Gorg proposed four policy implications regarding employment: Firstly that offshoring leads to higher job turnover in the short run. Secondly, low skilled workers suffer, higher skilled may benefit but no evidence of overall increased employment in the long run. Thirdly, different studies result in conflicting results; and fourthly, globalization leads to structural changes in advanced economies from manufacturing to service sectors. But causal relationships are still to be investigated.

Thirdly, the dynamic and contradictory nature of relationships associated with backshoring. The underlying reasons could be a mixture of changes in policy, costs, customer requirements, market and / or business strategic plans. Either when poor decisions are taken at an early stage, or when institutional pressures change so work may be returned (or backshored) to the home country. We need to better understand when backshoring is simply the consequence of an over enthusiastic initial response to the competition, a response to a radical change in the cost and business model or the more recent political and institutional pressure in the 'national interest'. Today, new institutional rules need to be defined that reflect the economic uncertainty and a lack of stability (Lane, 2008).

3   Data & Methodology

A mixed methods approach to a case study methodology is adopted with competitive comparisons drawn across airline and engineering sectors for both UK and German headquartered MNCs. Eight semi-structured interviews with nine senior executives in Germany, UK, India and Poland were undertaken for the research. Initial research questions were refined and additional data requested. Further interviews were undertaken with supplementary visits to host and supplier locations. (see Table 7). Opportunistic interviews took place in China and more are now planned to follow. The responses were analyzed on a qualitative basis and where secondary data has been obtained quantitative analysis has also been carried out in respect of comparative costs and performance. Data may also be triangulated by checking responses at different levels and in different parts of an organization. Initial interviews were with senior executives. Subsequent visits and interviews will be intended to include middle management, staff and where possible, trade unions / works council representatives. Because the case studies inevitably comprise different sections of a business rather than

the organization as a whole the 'unit of measure' will be important in making comparisons and drawing wider implications. The methodology can be summarized as:

Table 1. Selected Combination of Approaches (author adapted from Saunders et al)

CRITERIA	SELECTION
Philosophy	Pragmatism – combining positivism and interpretivism
Approach	A combination of deductive and inductive
Strategy	Multiple case studies that are paired by sector with multinational corporations MNCs who are significant market players. To support the case studies some additional secondary data and / or research of archive material will be required to triangulate the findings.
Choice	Mixed methods
Time horizon	Cross sectional with some historical perspective to current time
Techniques & Procedure	Semi structured interviews, recorded transcripts, analysis using a mixture of quantitative and qualitative techniques, supplemented with additional secondary data collection.

3.1   Developing the Conceptual Framework

It has been suggested that a firm's decisions might evolve from initial cost saving through the outsourcing of support activities as a first stage of disaggregating the value chain and then process improvement and further leveraging of labor cost savings through offshoring. Finally, if the economic circumstances in the home market change then politicians might in some manner influence MNCs to reverse their policy and restore work back into the home market – backshoring or similar (McKinsey, 2012). While this appears logical at a generic level, it may be rather too simplistic, especially at the level of a firm. Let us develop a more rigorous approach.

3.1.1   Proposed theoretical conceptual framework

A taxonomy of the relationships between LMEs and CMEs and their predicted approach to outsourcing and offshoring activity is shown below in Table 2. The first column distils the key questions that have been identified towards outsourcing and offshoring. Column 3 lists what are considered to be key dimensions to be explored through the research and subsequent analysis. Columns 4 and 5 represent hypotheses of anticipated responses if the companies conform to the stereotypical national LME model for the UK and CME for Germany.
It is intended that this conceptual framework and taxonomy will help in exploring case study differences in the rationale, success and lessons between the UK and Germany for each of the airline and engineering sectors as an empirical focus. The variables or dimensions chosen include the choice of location for outsourcing and / or off-

shoring which is essentially the reason or motivation that the company has for making the change, the control and coordination mechanisms in place, the levels of involvement and participation and finally, an ability to cope with changes in circumstances. The UK and Germany are compared using differing concepts of varieties of capital. The assumptions set out below and summarized in Table 2 are drawn from the literature (Lane, 1998; Lane and Probert, 2009; Whitley, 1997) in some cases reflecting a view that LMEs and CMEs are polar extremes, in other cases that over time there is some convergence and middle ground.

Taking each in turn, it is predicted that the motivation for outsourcing and offshoring will differ in that an LME will focus on short term cost cutting, budget control and shareholder interests. Initially, arbitrage of lower wages will be an inducement. If offshore they might also have a preference for English language speaking countries and traditional trading zones. On the other hand CMEs whilst also regarding low cost as a 'given' will focus on medium and longer term benefits in quality and performance and therefore a reluctance to outsource losing control and potentially intellectual property, if they offshore preferring central or European locations with a cultural or language similarity. This makes assumptions, such as all companies in a particular country will to at least some extent mirror and practice some of the characteristics associated with that classification of VoC. Also, the model can be regarded as rather static when in reality countries, sectors, markets and individual company approaches are dynamic and adapt to differing economic situations. So for countries such as Poland, Hungary or the Czech Republic the VoC positioning may be regarded by some as having shifted from a 'Transitional' positioning to a 'Pluralist Private Enterprise' (LME) or even to a 'Mixed' central position.

Thus there is a link to the second dimension of ownership and related aspects such as control and coordination and degrees of autonomy. This draws on GPN theory to the extent that policy and practice become embedded in the supply chain, the network and the territory. Also LMEs might be expected to be heavily focused upon the needs of the shareholder, strict cost and budget control as referred to above and an arm's length approach towards strategy – do what you have to do to meet budget and hence a high level of autonomy, as long as the local business stays within budget. A CME however, might be expected to be more likely to follow a multiple stakeholder model with a balanced approach to the differing needs of customers, suppliers, employees as well as shareholders; this is often referred to as market driven and customer focused. A CME might also be predicted to retain tight control over strategy, policy setting and resource allocation, and hence comparatively low levels of local autonomy, with a more hierarchical structure and somewhat slow to change with major decisions to be ratified centrally. A CME is therefore more constrained by institutional factors that influence managerial decisions such as 'what to offshore or outsource' and 'where to'?

The RBV and associated work on dynamic capabilities helps to inform us on how the lead company will manage core competences and resources. In deciding to transfer work from in-house and the home market are there than sufficient skilled resource to help the business transition work to either a third party or to an offshore subsidiary? With regard to managerial division of labor, LMEs might recruit local expertise with

only a minimum of expatriate managers. Such individuals are often attracted to the lifestyle and financial benefits and choose to stay longer term. In terms of cultural proximity they are more likely to be flexible and opportunistic with a low(er) level of concern other than an ability to speak and work in English where possible. CMEs may be predicted to invest more initially in setting up offshore operations with a comparatively high level of expatriate managers to transfer processes, set-up operations and organize training of a local workforce. Gradually they might transfer expertise to local management. Compared with LMEs a higher level of priority would be given to cultural proximity in terms of behaviors and language.

One of the key institutional factors to be explored is the role played by the trade unions and works council; and the inter-relationships with employees and management. For LMEs it is assumed that the influence is low or even non-existent, management will 'push the boundaries' once a decision has been taken within legal requirements and may be confrontational to enforce the decisions considered essential for the future of the business, especially at a time of poor economic prospects. CMEs on the other hand, will assume to be very much more consultative, actively avoiding confrontation and use times of growth to create jobs overseas and simultaneously move into key international markets.

Finally, we address evidence of a reversal in policy and returning work to the home country. For LMEs this might be influenced by political pressure or economic incentives. With CMEs we are assuming that this may be more likely to be a result of a change in market focus and /or strategy or a loss of intellectual property rights.

So, a theoretical projection is shown below in Table 2 presenting a series of hypothesis on what we might expect from a MNC headquartered in either the UK (LME) or Germany (CME). We have explored some relevant theory to underpin and construct this conceptual framework. The case studies will provide a 'test' for the conceptual framework of the theory both in use and practice. The first case study comparison is for airlines (UK and German) which will include passenger transport, cargo, maintenance and overhaul. The second case study is for engineering and manufacturing (UK and German) this covers products such as pumps, valves and seals for the offshore oil and gas industry together with software / hardware for the automotive components market. See Table 3 (airlines) and Table 4 (engineering) for summaries also further analysis in with preliminary findings (to date) in Tables 5-6.

Table 2   Conceptual Framework - Theoretical Projection

Question	Approach	Dimensions	Liberal market economy UK (LME)	Coordinated market economy GERMANY (CME)
What are the differences in the geographical, functional and temporal patterns of outsourcing and offshoring?	Outsource	Motivation	• Cost cutting and employee reduction • English speaking countries • Traditional trading zones	• Quality and performance, cost control is 'a given'. • Central / Eastern Europe preferred
How far do mechanisms such as ownership, control, coordination and the degree of autonomy differ?		Ownership	• Shareholder driven	• Multiple stakeholder
		Control & Coordination	• Arm's length on strategy. Strict cost and budget control	• Tight HQ control of strategy, policy and resources
		Degree of autonomy	• High – if meet financial targets then local control	• Low • Hierarchical structure • Can be slow to respond to change
How is this reflected in divergent international divisions of labor regarding the employment of indigenous or ex-pat managers?	Offshore or	Managerial Division of labor	• Low initial use of ex-pat managers who then stay on	• High initial use of ex-pat managers for set-up and training. Subsequently local management
To what extent do preferences for cultural proximity affect location?	outsourced offshore	Cultural Proximity	• Low, flexible, opportunistic	• High – language, behaviour

What is the influence of trade unions in the process of outsourcing and offshoring and how is this reflected in the structuring of the firms' labor markets?	or reverse offshore (Backshore)	• None, limited to legal requirements • Push the limits • Can be confrontational to enforce desired changes	• Consult widely • Actively avoid confrontation • Opportunistic – use growth to create additional jobs elsewhere
What evidence is there, and why of a reversal in policy – backshoring / reversed offshoring / outsourcing?	Change of policy	• Loss of initial cost-benefit. • Political pressure or economic incentives	• Loss of intellectual property • Change in market focus or strategy

4 Empirical Analysis

4.1 Discussion of the transport sector. (Let us call the UK airline 'A' and the German airline 'B').

With only initial semi structured interviews to date at both 'A' and 'B', albeit in some depth and detail, it is only possible to draw some general points regarding answers and relevance to the research questions. It is hoped to follow up these interviews shortly and also to visit some of the countries involved with the offshore activity. Apart from seeking a local perspective it would be appropriate to talk with a range of levels of personnel, not just senior executives. Trade Union and Works Council members are included and discussions are still on-going.

The two competitors selected in the airline sector show differences in approach. Both have moved back office support services and administration offshore, but the German organization has set up wholly owned shared service center's 'nearshore'; whilst the UK company moved processes to India, then as the business unit developed it was demerged and contracts are now in place to buy increasing levels of service back into 'A' from the offshore and outsourced provider. With engineering, repair and mainte-nance work, also catering the approaches are again different. The German company 'B' retains control and manages cost by leveraging labour costs offshore and using agency employees where necessary although this can cause questions around control. The UK business however works through its procurement and contracts team to place work either offshore or outsourced or both to keep costs down. 'A' have now learnt to manage these contracts more effectively and even buy in catering and engineering services from the competitor 'B' when appropriate in best value terms. Where labor costs are less of a concern they have improved processes now to such an extent they are prepared to reverse a previous policy and bring work back into 'A' where it now cheaper following efficiency savings. 'A' aim for flexibility and an ability to react to market changes. The yield and volume of seat tickets sold are carefully monitored with metrics such as unit costs for an available seat per km. With price reductions and discount promotions, again the cost base is carefully monitored (with and without fuel costs that cannot be controlled). Productivity improvements have to fund pay awards; efficiency improvements are regarded as important with large volume activity.

For a summary of findings and comparison with conceptual framework (see Appendix Table 3). The key challenges for the Airlines include (see also Appendix Table 4):

1. Highly competitive, overlapping segments in the market e.g. low cost passenger travel, and price competition for larger organizations.
2. Network of partner and alliance companies for global coverage.
3. Passenger transport and engineering businesses can be counter cyclical.
4. Profitability is sensitive to fuel costs, economic conditions and competition.
5. Customer loyalty is a key factor in a high profile customer service business.

6.   Differing levels of power, control and influence between management, trade unions and works council.

4.2   Discussion of the manufacturing and engineering sector. (Let us call the UK engineering company 'C' and the German engineering company 'D').

Once more with only initial semi structured interviews to date at both 'C' and 'D', albeit in some depth and detail it is only possible to draw some general points regarding answers and relevance to the research questions. It is hoped to follow up these interviews shortly and also to visit some of the countries involved with the offshore activity. These two engineering case studies do therefore provide some early insights on differences in approach with respect to competences, technology transfer around the world and the development of key alliances; as postulated by Lynn and Salzman (2009).

There are similarities in focus for both UK and German companies – to initially cut costs, keep prices down and then to improve efficiencies, processes and customers service. The method of delivery however, is different. The UK company 'C' takes a long term view but with short term deliberate steps towards partnership and then integration and acquisition utilizing outsourcing and offshoring where appropriate. The German company 'D' however, prefers to retain centralized control by establishing a subsidiary business offshore from the outset, with no or little consideration of outsourcing. There is also little evidence of synergies across the German group. Both 'C' and 'D' companies have grown and employment has been largely protected, although the United States division of 'C' has reversed a policy to move work to Mexico back into the US. It would also seem that complex work offshored to India by 'D' has subsequently had to be re-worked in India.

For summary of findings and comparison with conceptual framework (see Appendix Table 5). The key challenges for the engineering businesses include (also see Appendix Table 6):

1.   On-going cost control, especially in the UK company which is Shareholder driven.
2.   Customers ask for, and expect lower prices and local supply.
3.   Competitor pressure within the market and industry sector.
4.   Preferred tendency with 'C' to try a joint venture and then acquisition, integrate and restructure to reap rewards.
5.   More control if it is a wholly owned subsidiary of 'D', can then avoid issues of IP with a third party.

5   Conclusions

It is well known that German has managed its economy in such a way that it has been less exposed to the economic pressures suffered by much of the rest of Europe.

To some extent this has allowed management to move operations offshore but not outsource, gain the benefit of lower costs (some 10 per cent at least) without losing jobs at home. However, as costs increase at a faster rate in many overseas markets the search for productivity benefits and efficiency gains continues. The basic components of a 'coordinated market economy' seem to prevail with evidence of institutional coordination, long term planning but also central control and an aversion to risk. The UK companies in both case studies were quicker to outsource, favored short term cost savings but were also more flexible and agile, taking risks with trade unions and suppliers and customers to seemingly favor shareholders. In many respects this is consistent with the 'liberal market economy' capitalist model. In both cases the choice of location was often different, as was the approach to delegation and autonomy suggesting differing views on governance. The underlying theoretical constructs of varieties of capitalism, the resource based view and global production networks were each found to be of value. (Research Questions 1 & 2, Tables 3 & 4).

German Companies use expatriate managers for the short term but then mostly rely on local skills. UK companies use local staff from the outset. German companies also place more emphasis on language, near shoring and cultural empathy (Research Questions 3 & 4 , Tables 3 & 4). UK companies may have a tendency to be adversarial with trade unions, forcing job reductions when considered to be essential whereas German companies were cooperative and averse to conflict where possible. (Research question 5, Tables 3 & 4). Only isolated cases of reverse offshoring or backshoring were evident from the four companies. (Research question 6, Tables 3 & 6).

References

Banker, R.D. et al (1984) Estimating the most productive scale size using data envelopment analysis. *European Journal of Operational Research*. July, 35-44.

Barney, J. 1991. Firm resources and sustained competitive advantage. *Journal of Management,* 17**,** 99-120.

Barthélemy, J. (2011), Editor's Choice, feature on Offshoring*, Financial Times*, May 26, 2011.

Coe, N.M., Hess, M. et al (2004) Globalizing regional development: a global production networks perspective *Trans Inst British Geographers* 29 (4): 468-484.

Contractor, F.J. et al (2010). Reconceptualizing the Firm in a World of Outsourcing and Offshoring: The Organizational and Geographical Relocation of High-Value Company Functions, *Journal of Management Studies* 47:8 December.

Contractor, F.J. (2012). Why do Multinational Firms exist? A theory note about the effect of multinational expansion on performance and recent methodological critiques. *Global Strategy Journal* 2, 318-331, Wiley.

Dunning, J. H. (1988). E*xplaining International Production*. London, Unwin Hyman.

Economist (2011), Business, multinational manufacturers, *Economist,* May 14, 2011, pg 73.

Gorg, H. (2011). *'Globalization, offshoring and jobs'*. ILO / WTO report.

Granovetter, M. (1992) *'*Economic institutions as social constructions: a framework for analysis'*, Acta Sociologica*, 35, 3-11.

Hall, P. and Soskice, D. (2001).*' Varieties of Capitalism: The Institutional Foundation of Comparative Advantage'*. Oxford, Oxford University Press.

Harms, P. Lorz.O. et al (2009). "*Offshoring along the Production Chain",* Study Center Gerzensee Working Paper No. 09.01 and CESifo Discussion Paper No. 2564.

Hess, M. and Wai-Chung Yeung (2006), *'*Whither Global Production Networks in economic Geography? Past, Present and Future', *Environment and Planning,* Vol. 38

Kirkegaard, J.F. (2008). Offshoring, Outsourcing and Production Relocations—Labor Market Effects In The OECD and Developing Asia, *The Singapore Economic Review,* Vol. 53, No. 3, 371–418

Kohler, W. (2008). *Offshoring: Why do stories differ?* CESifo Working Paper 2231.

Lane, C. (1998).' European Companies between Globalization and Localisation: A Comparison of Internationalisation Strategies of British and German MNCs', *Economy and Society*, 27(4), 462-485

Lane, C. (2008). 'National Capitalisms and global production networks: an analysis of their interaction in two global industries'. *Socio-Economic Review*. Vol (6) pp 227-260.

Lane, C. and Probert, J. (2009).' *National Capitalisms, Global Production Networks Fashioning the Value Chain in the UK, USA and Germany'*. Oxford, Oxford University Press.

Leibl,P., Morefield, R. et al. (2010). *A study of the effects of backshoring in the EU. Proceedings of the 13th International Conference.* American Society of Business and Behavioural Sciences.

Lynn, L. and Salzman, H. (2009). 'The new globalization of Engineering: How offshoring of advanced engineering affects competitiveness and development'. *Economics, Management and Financial Markets*, 4, 1. March. New York, Addleton.

McKinsey (2012). Manufacturing the future: The next era of global growth and innovation. *McKinsey Global Institute and McKinsey Operations Practice*, November.

Mitchell, A.J. and Hill, S. (2004) The Great M&A Gamble, *The Ashridge Journal 360°*, Autumn.

Morgan G. et al (2006). *'Changing Capitalisms? Complementarities, Contradictions and Capability Development in an International Context',* pp78-109, Oxford, Oxford University Press.

O'Connor, S. (2013). *'Data show yawning productivity gap between UK and G7 peers'. Financial Times.* September 18th 2013.

Oliver, C. (1997) 'Sustainable competitive advantage: combining institutional and resource based views'. *Strategic Management Journal,* 18:9, pp. 697-713.

Pisano, G. (2009)' Restoring American Competitiveness', *Harvard Business Review,* July-August.

Soulsby A. and Clark E. (2007). 'Organizational theory and the post socialist transformation: Contributions to organizational knowledge.' *Human Relations* Vol. 60(10), Sage.

Whitley, R. and Kristensen, P. H. (eds) (1997). *'Governance At Work: The Social Regulation of Economic Relations'.* Oxford, Oxford University Press.

APPENDICES

Airline case summary

Table 3   UK and German Airlines Compared

Question	Approach	Dimensions	Liberal market economy UK (LME)	Coordinated market economy GERMANY(CME)
What are the differences in the geographical, functional and temporal patterns of outsourcing and offshoring?	Outsource	Motivation	India, South Wales. Cost and reduced employees numbers. Catering, administrative and revenue accounting, engineering, maintenance, repair and overhaul.	Poland, China, Thailand, Mexico. Quality, performance and cost. Shared services, ticket booking, invoicing, maintenance, repair and overhaul.
How far do mechanisms such as ownership, control, coordination and the degree of autonomy differ?		Ownership	Outsource: Shareholder value	Retained offshore subsidiary
		Control & Coordination	Offshore and outsourced. Arm's length, market driven. Open book, service level agreements. Procurement led/ contract driven.	Tight HQ organizational control

Question		Category		
How is this reflected in divergent international divisions of labor regarding the employment of indigenous or ex-pat managers?	Offshore or	Degree of autonomy	Generally high. Maintenance retained at an internal subsidiary.	Low, but Increasing, based off-shore or nearshore
		Managerial Division of labour	Local staffs. No ex-pats.	Run by ex HQ managers. At start-up managerial level withdraw at operative level as soon as possible and recruit locals
To what extent do preferences for cultural proximity affect location?	outsourced offshore or	Cultural Proximity	Unimportant. Global reach.	Important – language & culture. Focus on regions Europe, SE Asia, S America.
What is the influence of trade unions in the process of outsourcing and offshoring and how is this reflected in the structuring of the firms' labor markets?	reverse off-shore (Back-shore)	Relationship with employees / Trade Unions	Adversarial, non co-operative	Cooperative, aversion to conflict
What evidence is there and why, of a reversal in policy – backshoring / reversed offshoring / outsourcing?		Change of policy	MRO work retained / returned in-house	Not so far

Table 4: Interpretation of initial pilot research questions – headline comparison of Airline sector case studies - approaches to offshoring and outsourcing. Source: author

RESEARCH QUESTION	FINDING
To what extent are German and UK multinational companies displaying different varieties of capitalism and how does that effect decisions and strategies related to the deployment of outsourcing and offshoring?	Both organizations have responded to an industry need to cut costs, responding to low cost competition and customer needs. Both have deployed a mixture of outsourcing and offshoring tactics. In Germany these have developed and grown as a contributor to long term strategy, and the need for tight governance that has been acknowledged by a new CEO. Control is exercised through keeping units as subsidiary businesses within the group. The UK model is not purely cost driven, there is a determination to seek value for money and reverse earlier decisions by in-shoring or in-sourcing when productivity improvements have been realized.
What is distinctive about the governance of German and UK multinational firms?	A sense that for the German firm power and key decisions are centralized and taken at a high level in the organization. Added value often assumed to come from head office in Germany. For the UK business there is more evidence that the Procurement dept. have considerable power for contracts and post-acquisition integration in conjunction with legal. Work is generally put out to competitive tender on a periodic basis any exception must then be approved by the UK main board.
How is the above reflected in idiosyncratic patterns of outsourcing and offshoring at both a national and sector level?	For Germany, a very careful selection of preferred countries, with a reluctance to engage with India because of language and cultural differences. Search for cities where similar work is clustered e.g. Krakow. For the UK a willingness to use the German competitor (in the core business) where there is a clear business case and conflict of interest.

Which functions or processes are moved offshore, where to and why?	Non-core, administrative processes e.g. revenue accounting, catering, engineering and maintenance.
In what ways does the embeddedness of firms influence the motives, control and strategy of the parent multinational company?	The German company has concentrated shared service centers at several locations around the world to serve key global markets. As they grow so the significance increases of these sites in Krakow, Bangkok and Mexico. As part of the restructuring of the new UK holding company board, responsibility is retained for overall strategic direction at group level.
To what extent are outsourcing and offshoring policies reversible, and what is the experience in Germany and the UK?	To date there seems to have been no evidence of reversing policies with the German business. However, the UK business have reversed and insourced work, previously given to third party contractors following internal productivity improvements, that both freed capacity and rendered the operations cost effective.

Engineering Case Summary

Table 5   UK and German Engineering compared

Question	Approach	Dimensions	Liberal market economy UK (LME)	Coordinated market economy GERMANY (CME)
What are the differences in the geographical, functional and temporal patterns of outsourcing and offshoring?	Outsource	Motivation	UK, Czech republic, China Less keen on India Catering, administrative and revenue accounting, engineering, maintenance, repair and overhaul. Cost	India, Vietnam, Czech Republic – 'lead' global roles in Asia, Europe and North / South America. Embedded software applications, IT systems, accounting, call centers. In Czech Republic – the development of new automotive platforms; R&D, Engineering and Manufacturing. Local expertise and cost.
How far do mechanisms such as ownership, control, coordination and the degree of autonomy differ?		Ownership	Offshore through Joint Venture then wholly owned acquisition. Financial control via HQ, but freedom to run business locally.	Now wholly owned, offshore subsidiaries, budget control and OEM contact through HQ.

				Global operations via HQ	HQ with OEM, divisional control and global coordination from HQ
Control & Coordination					
Degree of autonomy				Relatively high	Relatively high in terms of design and delivery. Close budget and resource planning and monitoring from HQ.
Managerial Division of labor	How is this reflected in divergent international divisions of labor regarding the employment of indigenous or ex-pat managers?	Offshore or	outsourced offshore	Kept to a minimum	Ex-pat initially as senior manager. Replaced with local after 5 years, maybe 5 ex pats out of 10,000 local employees. In Czech Republic initial training of engineers in Germany then on-site over 2 years. Ex pats may stay.
Cultural Proximity	To what extent do preferences for cultural proximity affect location?	or		Significant preferences through experience	Less important – although with the Czech Republic there are advantages of proximity, similar markets, some

		reverse off-shore (Back-shore)	
			ease of language and cultural affinity.
What is the influence of trade unions in the process of outsourcing and offshoring and how is this reflected in the structuring of the firms' labor markets?	Relationship with employees / Trade Unions	Redundancies where required	Avoid conflict, timed to coincide with growth to avoid job losses in Germany. Few issues in Czech republic – weak union but also free labour market and plant growth offering security.
What evidence is there, and why, of a reversal in policy – backshoring / reversed offshoring / outsourcing?	Change of policy	Mexico back to the US	Stories of complex work being returned from India to Germany for rework.

Table 6: Interpretation of initial pilot research questions – headline comparison of Engineering sector case studies: approaches to offshoring and outsourcing. Source: author

RESEARCH QUESTION	FINDING
To what extent are German and UK multinational companies displaying different varieties of capitalism and how does that effect decisions and strategies related to the deployment of outsourcing and offshoring?	To some extent the differences here are subtle rather than significant. There is some evidence of Corporatist / coordinated behavior in Germany and liberal by the UK company. A huge reluctance to outsource anything other than Travel services by the German organization is apparent. The similarities are common – both employ high quality engineers and other specialists, both are keen to cut cost and improve efficiencies. Both have grown and are successful.
What is distinctive about the governance of German and UK multinational firms?	The role of MNC in transferring technology is a key FDI flow. Both cases meet the usual criteria high R&D, large share of professional and technical workers, complex technical products, high levels of differentiation. Advantages come from Ownership, location and internalization (Dunning, 1988); and democratic countries such as India and Czech Republic tend to attract more FDI with lower country risk, debt risk. The relationship between headquarters and subsidiaries can be understood by Agency Theory and ordinarily competition would encourage coordination and cooperation, between and across subsidiaries. What is unusual with the German case here is that there is little communication across the group only between headquarters and a specific subsidiary.
How is the above reflected in idiosyncratic patterns of outsourcing and offshoring at both a national and sector level?	The UK case suggests that they will deploy whatever approach is most applicable, especially for short term gain; also that the German organization will avoid outsourcing in favor of controlled offshoring.

43

Which functions or processes are moved offshore, where to and why?	Not so much functions as products and then the entire business support system that is required for those products in both Germany and the UK.
In what ways does the embeddedness of firms influence the motives, control and strategy of the parent multinational company?	In Germany long term development of FDI has resulted in considerable growth and recognition that maturity is now close to optimum in India leading to the establishment of a second, smaller clone in Vietnam. The UK company have restructured and developed a global strategy, a current priority of which is to coordinate common IT platforms across the sites.
To what extent are outsourcing and offshoring policies reversible, and what is the experience in Germany and the UK?	None observed here within Europe but the US division of the UK Engineering company has reversed a policy to move work from the US to Mexico.

Table 7: List of semi structured face to face and telephone Interviews

(typically 1hour each)

Case Study Airlines		Case study Engineering	
UK Company 'A'	**German Company 'B'**	**UK Company 'C'**	**German Company 'D'**
London Nov 2011 Procurement / strategy manager	Cologne Oct 2011 VP services	Slough Dec 2011 VP Ops	Stuttgart Oct 2011 VP Eng
Follow-up Nov 2012	Hamburg Oct 2011 Director Engine Lease	Slough Dec 2011 Director Group Ops	India Dec 2011 Company President
	Krakow Jan 2012 MD Service Centre		Prague Jan 2013 Director
	Follow-up Nov 2012	Follow up Nov 2012	Follow up Nov 2012
Trade Union Nov 2013 Academic, former TU National Officer			

A System Dynamics Framework for Water Footprint Management in Agrifood Supply Chains

Eirini Aivazidou [1], Naoum Tsolakis [1], and Eleftherios Iakovou [1]

[1] Laboratory of Quantitative Analysis and Advanced Supply Chain Management, Department of Mechanical Engineering, Aristotle University of Thessaloniki
aveirini@auth.gr

Abstract. Nowadays, a prominent concern is the overexploitation and scarcity of the global freshwater natural resources, which have led to a plethora of negative social, environmental and economic impacts. To that end, the concept of water footprint has been introduced at a national, corporate, as well as product level. Specifically, companies map their water footprint in order to assess the sustainability of their production processes or their entire supply chain networks. This is of particular interest for the agrifood industry, since the major part of the national water footprints stems from the agricultural production and food processing. In this context, we propose the employment of System Dynamics, a simulation-based modeling technique, for capturing the dynamics that either promote or limit water sustainability in the agrifood sector in order to foster water footprint management. Specifically, in this manuscript we intent to: (i) first present the water footprint of the agricultural and industrial production in the EU-15 countries, and (ii) then propose a first-effort generic conceptual System Dynamics framework that could assist governmental authorities, regulators and business managers in the design and implementation of effective policies and strategies for water footprint management across agrifood supply chains. Since System Dynamics modelling for water footprint management is rather limited in the existing literature, the development of the proposed methodology could provide important managerial insights about national policy-making and corporate decision-making interventions towards sustainable development in the European agrifood sector.

Keywords: Agrifood Supply Chain, Policy-making, Sustainability, System Dynamics, Water Footprint Management

1   Introduction

A global phenomenon that raises major challenges for policy-makers and businesses nowadays is the rapidly depleting freshwater supply in many regions of the world, which leads to water scarcity (UNESCO, 2006). Freshwater resources are directly threatened by human activities and expected to be further affected by anthropogenic climate change (Herath et al., 2013). Specifically, population growth, economic de-

velopment and industrial pollution intensify the stress over global freshwater supplies (Rockstrom et al., 2009). To that end, the increasing freshwater depletion has put high pressure on the globalized economy and thus balancing the economic, environmental and societal impacts of global water consumption has emerged as a key issue for governments, companies and consumers (McKinsey, 2009).

In this context, researchers have introduced the concept of water footprint (WF), a multidimensional indicator for assessing both direct and indirect water consumption and pollution (Hoekstra et al., 2011). The magnitude of WF differs among various economic sectors and production systems. Agrifood industry is considered as one of the most water-consuming industries worldwide. Indicatively, the agricultural sector utilizes 70% of the global water resources (Aviso et al., 2011), while the industrial sector is responsible for 22% of the global water consumption (WBCSD, 2006). To that effect, managing WF across agrifood supply chain (AFSC) networks is imperative for the establishment of economic, environmental and societal sustainability within the agrifood sector.

Freshwater is a pivotal constituent of the supply chain operations for the majority of companies (Ercin et al., 2012). Empirical evidence documents that proper management of freshwater supplies relates to multiple benefits for businesses, namely: (i) building of a green corporate image, (ii) compliance with increased regulatory measures, (iii) control of financial risks related to pollution and environmental damage, and (iv) stability in business operations due to water sufficiency in the production/manufacturing processes (Pegram et al., 2009; Rondinelli and Berry, 2000). Therefore, quantifying and mapping WFs has been an evolving field of study since the introduction of the WF concept (Hoekstra, 2003; Hoekstra and Mekonnen, 2012). However, since there is still not any legislative scheme that improves WF management in corporations, policy-makers need to assess the total national freshwater consumption during production processes in order to have an indication for the establishment of relevant regulatory interventions.

In this manuscript, we investigate the role of WF management in supply chains in the European region and develop a conceptual methodological framework in order to foster water sustainability in the agrifood sector. Therefore, the objective of this research paper is twofold: (i) to present the WF of the agrifood and industrial production of the EU-15 Member States, and (ii) to propose a first-effort generic System Dynamics (SD) framework that could assist governments, regulators and businesses in designing and implementing effective policies and strategies for WF management across AFSCs.

The remainder of the paper is structured as follows. In Section 2, we present the research background regarding WF definition and methodology, as well as recorded estimations of the total WF on the EU-15 level, while we also describe the typical structure of an AFSC. Following, Section 3 includes: (i) a brief introduction to the SD methodology, (ii) a description of the system under study, along with the related variables and the main assumptions of the modelling approach, and (iii) the development of the proposed SD model for the WF management across an AFSC. Finally, Section 4 sums-up with the conclusions and the recommendations for future research.

2 Research Background

In this section, the research background of the proposed WF conceptual framework is provided. More specifically, we introduce the concept of WF, while we provide real-world data about WF volumes in the EU-15 countries regarding the economic sectors of agricultural and industrial production, as well as domestic water supply. We also present the structure of a typical AFSC in order to highlight the importance of WF management in the agrifood sector.

Water Footprint Definition and Methodology

Water is a critical element of production systems and the base of the ecosystem (Yu and Qingshan, 2014). Thus, during the last decade the impact of human consumption on water resources was mapped through the concept of WF assessment.

The term WF was initially introduced by Hoekstra and Hung (2002), as a measure of a nation's total consumption of freshwater resources used for products which are consumed within the country. Specifically, the WF of national production is defined as: "*the total freshwater volume consumed or polluted within the territory of the nation as a result of activities within the different sectors of the economy; it can be calculated by summing the WFs of all water consuming or polluting processes taking place in the nation*" (Hoekstra and Mekonnen, 2012). Further, corporate WF is defined as the total volume of freshwater used directly and indirectly to run and support a business. The business WF consists of two components: (i) the direct freshwater consumed or polluted by the producer during business operations (i.e. production/manufacturing or supporting activities), and (ii) the indirect freshwater used across the rest supply chain network (Ercin et al., 2011; Hoekstra, 2008). At a product level, WF is considered as the total volume of freshwater used directly or indirectly to produce a good or a service, which equals to the sum of water consumed and polluted across the product's full supply chain (Hoekstra, 2008).

WF is a multidimensional indicator, which refers not only to the volume of freshwater consumed and polluted but also to the type of water used, as well as to the temporal and spatial dimension of the water consumption (Hoekstra et al., 2011). In this context, WF is classified into three (3) specific types, namely blue, green and grey. According to Hoekstra et al. (2011), blue WF is the volume of surface or groundwater consumed during the production of a good or a service. Consumption refers to the volume of freshwater: (i) evaporated, (ii) incorporated into the product, (iii) abstracted from a catchment and then return to another catchment or the sea, and (iv) withdrawn from a catchment in one time period and return in another period. Green WF constitutes the volume of rainwater consumed during the production process, and especially in the farming stage of the agricultural products (Hoekstra et al., 2011). Consumption refers to the volume of precipitation water: (i) evaporated from the fields, (ii) transpired through plants, and (iii) incorporated into the harvested crop or wood. Finally, grey WF is considered as an indicator of freshwater pollution associated with the production chain of a product. It is the volume of freshwater required to assimilate the load of waste, which is quantified by the volume of water required to dilute pollutants

to such an extent that water quality remains above the agreed quality standards (Hoekstra et al., 2011).

In this context, a methodology for Water Footprint Assessment (WFA) was developed in order to foster the protection and management of freshwater resources (Hoekstra et al., 2011). According to Hoekstra et al. (2011), the WFA approach is comprised of four (4) distinct phases, namely: (i) setting goals and scope, (ii) WF accounting, (iii) WF sustainability assessment, and (iv) WF response formulation. Particularly, quantitative indicators are utilized in order to map freshwater consumption and pollution during the accounting phase, while a multifaceted analysis regarding multiple environmental, economic and social aspects of water use is performed during the sustainability assessment phase (Hoekstra et al., 2011).

During the last years, there was a controversy concerning WFA and Life Cycle Assessment (LCA) with regard to water use (Hoekstra et al., 2009). However, the two methodologies exhibit specific strengths and benefits. According to Boulay et al. (2013), both WFA and LCA utilize quantitative indicators yet in different stages of the assessment. Specifically, WFA quantifies the volume of water used (i.e. green, blue, grey WF), while LCA focuses on the quantification of the environmental impacts from depriving ecosystems of freshwater resources (Boulay et al., 2013). At the same time, the core difference between the two methods is that LCA is mainly product-driven, while WFA is a more comprehensive water management-driven approach that tackles the analysis of the sustainable, efficient and equitable allocation of freshwater resources with either a national, a corporate or a product focus (Boulay et al., 2013).

EU-15 Water Footprint Statistics

Mapping the total WF of a country is crucial for developing national policies and strategies with pivotal implications for the three pillars of sustainable development: environment, society and economy. According to the up-to-date research, one can distinguish three (3) main water-using production sectors: (i) the agricultural sector, (ii) the industrial sector, and (iii) the domestic water supply sector. To that effect, we have gathered data about the WF regarding in these sectors of the EU-15 countries within the decade 1996-2005. This period was selected due to the availability of relevant data (retrieved from WaterStat, a WF Network's global water database, www.waterfootprint.org).

Fig. 1 illustrates the global annual average WF (in Mm^3/yr) of the EU-15 Member States. The WF related to agricultural production (crop production, grazing and animal water supply) takes the largest share in the total WF within EU-15 countries. France, Spain, Italy and Germany are the countries with the largest green WF within their territories: 68,372 Mm^3/yr, 57,091 Mm^3/yr, 48,449 Mm^3/yr and 40,409 Mm^3/yr respectively, which constitute 23.1%, 19.3%, 16.4% and 13.6% of the total EU-15 green WF. Spain also has the largest blue WF (15,695 Mm^3/yr), which corresponds to 5.3% of the total EU-15 blue WF. Moreover, regarding the grey WF, France (15,894 Mm^3/yr), Italy (15,222 Mm^3/yr), and Germany (14,189 Mm^3/yr) have the largest amounts of the total EU-15 gray WF (5.4%, 5.1%, and 4.8% respectively).
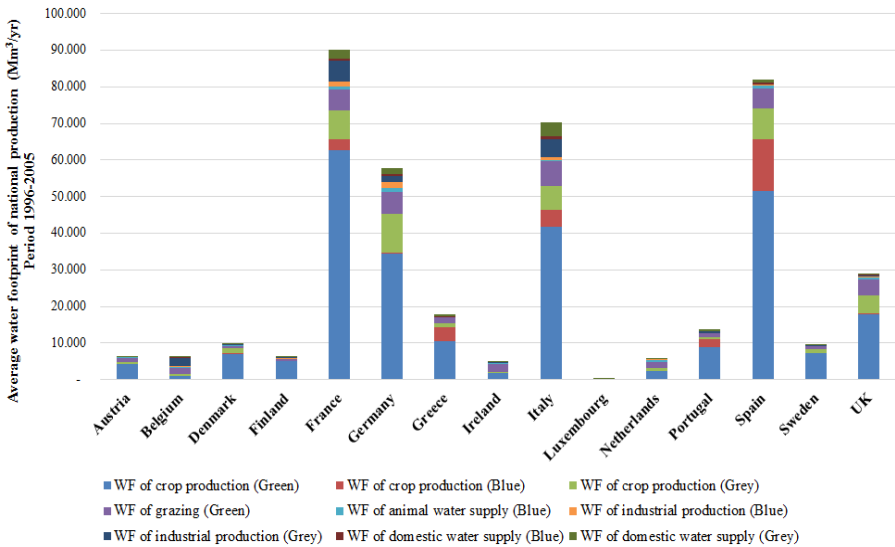
Fig. 1. Average WF of national production (Mm^3/yr) (Period 1996-2005).

Due to comparison reasons, China, India and the United States are the countries with the largest total WFs with an average of 1,207,393 Mm^3/yr, 1,181,663 Mm^3/yr and 1,053,462 Mm^3/yr respectively (Hoekstra and Mekonnen, 2012). To that effect, these three countries account for 38% of the WF of production worldwide.

Agrifood Supply Chain Description

During the last years, the agrifood industry has started embracing supply chain management as a strategic concept for its competitiveness (Tsolakis et al., 2014). According to Tsolakis et al. (2014), the wide range of activities across an AFSC, such as farming, storage, transportation and food production/processing, can have significant impacts on environmental sustainability. In addition, consumers are highly aware of environmental sustainability issues and therefore the application of sustainability practices can create potential bases for competitive advantage for all supply chain members, especially in the food sector (Bourlakis et al., 2014; Carter and Dresner, 2001).

In general, an AFSC covers a variety of activities in a "farm-to-fork" sequence, including farming, industrial processing, packaging, transportation, warehousing, as well as distribution and retailing (Iakovou et al., 2012). As shown in Fig. 2, there are five flow types among these operational echelons, namely: (i) physical material and product flows, (ii) process flows, (iii) financial flows, (iv) information flows, as well as (v) energy and natural resources flows, which interconnect suppliers and consumers (Iakovou et al., 2012).
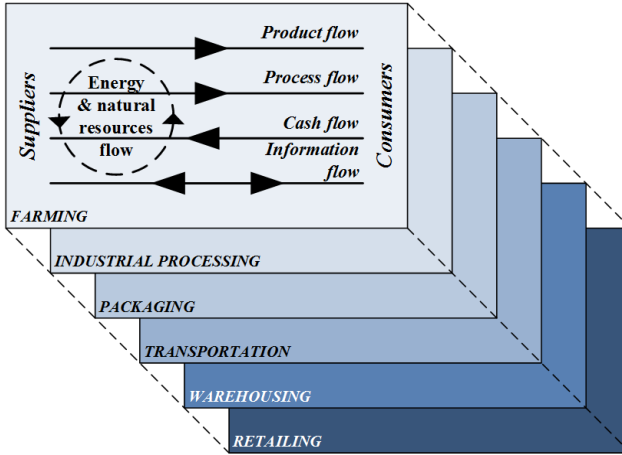
Fig. 2. AFSC echelons (Adapted from Iakovou et al., 2012).

Production efficiency and supply chain performance accounts for majority of the total WF in agrifood industries (Manzardo et al., 2014). Concerning the farming stage, agricultural production accounts for 92% of the freshwater resources of humanity (Gerbens-Leenes et al., 2013), while the global WF of animal production constitutes almost one third of the total WF of agricultural production (Hoekstra and Mekonnen, 2012), and this fraction is likely to increase (Liu et al., 2008). Considerable water consumption is also reported at other food supply chain stages, such as processing, logistics, packaging and distribution (Andersson et al., 1998; Ruini et al., 2013). Additionally, strict requirements for product quality and the associated hygiene issues in manufacturing contribute to large amounts of high-quality water consumed during food processing (Klemeš et al., 2009).

3   Modelling the Agrifood Supply Chain Water Footprint

In this section the SD methodology is introduced, while the SD modelling framework for WF management is analysed. This framework could be employed by decision-makers as a useful tool for assessing the impact of various intervention policies on promoting water sustainability across AFSCs.

System Dynamics Methodology

SD is a simulation methodology, which is allows the analysis and understanding of the behavior of complex systems over time. Originally, the tool was described and used by Forrester (1961) to examine the instable employment environment in General Electric, while over the years the SD tool was employed to conduct simulations of diverse scientific and engineering systems (Forrester, 1969). The feedback control characteristics of the SD approach render it the proper methodological tool for deci-

sion-making regarding a wide range of problems, from managerial and socioeconomic to organizational ones (Roberts, 1978; Sterman, 2000).

The core concepts of SD are feedbacks, causal loop diagrams, as well as stock and flow maps (Sterman, 2000). Feedback loops and structures are necessary because systems rarely exhibit linear behaviour due to the interactions among the physical and institutional configuration of the systems. Feedback structures are fundamental because they capture the real patterns or modes of a system's behavior, as they dynamically evolve through the time. Causal loop diagrams are used in order to capture the mental models of a system. Additionally, causal loop diagrams assist modelers in representing the feedback structures of a system. In a causal diagram, arrows describe the causal influences among the variables of the system (Sterman, 2000). A polarity is assigned to each arrow that indicates the relation between dependent and independent variables. A positive (+) polarity denotes that the effect changes towards the same direction as the cause (reinforcing feedback). On the other hand, a negative (-) polarity denotes that the effect changes towards the opposite direction of the cause (balancing feedback) (Sterman, 2000).

Water Footprint System Dynamics Model

SD applications have been quite extensively used in the past for water management (Winz et al., 2009). Indicative water resources management issues that have been tackled, including SD models on: water saving, wastewater reuse, and water transfer (Zhang et al., 2009), water conservation policy analysis (Ahmad and Prashar, 2010; Qaiser et al., 2011), community-based water planning (Tidwell et al., 2004), water trading/leasing and transfer schemes (Gastelum et al., 2010), irrigation water management (Khan et al., 2009), flood management (Ahmad and Simonovic, 2006; Simonovic and Ahmad, 2005) and reservoir operations (Ahmad and Simonovic, 2000).

However, to the best of our knowledge, there is a lack of SD studies that address WF management in supply chains, and especially of the agrifood sector. Therefore, we developed a first-effort conceptual WF model across AFSCs for the assessment of the impact of several interventionary policies associated with the sustainable WF management in AFSCs. Specifically, the nexus of physical and information flows that connect the green, blue and gray WF elements with the basic operations of an AFSC is presented.

System under study. To begin with, we assume that the proposed system includes the main echelons of an AFSC, namely farming, industrial processing, packaging, as well as warehousing and retailing that are considered as a single supply chain stage. In addition, we suppose that each echelon contributes to the three components of the total WF (i.e. green, blue, grey) according to the relevant supply chain operations.

Following, each AFSC echelon is represented by a stock variable which indicates the quantity of food product units produced at a specific point in time. Furthermore, green, blue and grey WFs are represented by stock variables that are quantified by the volume of water resources consumed or polluted at a specific point in time. Stock variables are increased or decreased by the relevant flow variables (rates). Since the lead times of the food transportation flows among the various supply chain echelons

are negligible compared to the simulation step, delays due to transportation are not taken into consideration. Additional control (i.e. green WF optimal level, etc.), and other variables that follow a known distribution (i.e. consumers demand, crop area availability, etc.) are considered, rendering the model more realistic.

Policies under study. Five (5) different policies, which are represented by decision variables and quantified by the relevant rates of the water resources savings, could help policy-makers to intervene in the WF system. These indicative policies are included in a set of sustainable practices for WF management that can be deployed in AFSCs (Aivazidou et al., 2014). Particularly, the policies under study are:

- Policy 1: Implementation of methods for water retention in the soil
- Policy 2: Implementation of precision irrigation methods
- Policy 3: Investments in water-efficient techniques
- Policy 4: Prudent use of pesticides/fertilizers
- Policy 5: Reduction of toxic packaging materials.

Causal loop diagram. The associated causal loop diagram illustrates the correlation among the variables that specify the behaviour of the reviewed conceptual WF system (Fig. 3). The system includes eleven (11) feedback loops. Indicatively, we present four (4) typical loops that define the SD model under study:

- Balancing loop B1: The amount of yield production in the farming stage decreases when the food processing rate increases. At the same time, an increase in the yield production raises the processing rate.
- Balancing loop B2: By definition the green WF increases as the rate of the water evaporation from the fields rises (this rate is also positively correlated to the average crop yield). Then, the real-to-optimal green WF ratio also increases when the level of total green WF rises (but decreases when the relevant control variable increases). The aforementioned ratio necessitates Policy 1 regarding the implementation of water retention methods, which finally reduces the rate of the water evaporation from the fields.
- Reinforcing loop R1: The amount of products in the retail stores augments the consumers' sales rate, which in turn increases the average consumer's sales. The latter also raises the retailers' sales rate, which in turn boosts the amount of the final products.
- Reinforcing loop R2: The yield production rate increases the production in the farming stage, which in turn correlates positively to the food processing rate. Then, the processing rate raises the amount of food production in the industrial processing stage. As the food production increases, the packaging rate and following the amount of packaged product units increase, thus fostering the retailers' sales rate. Hence, there are more final products in the warehousing and retailing stages, which increase the average food losses due to deterioration and spoilage. The increased food waste highlights the need to have a greater yield production rate so as to avoid any food shortage in the market.
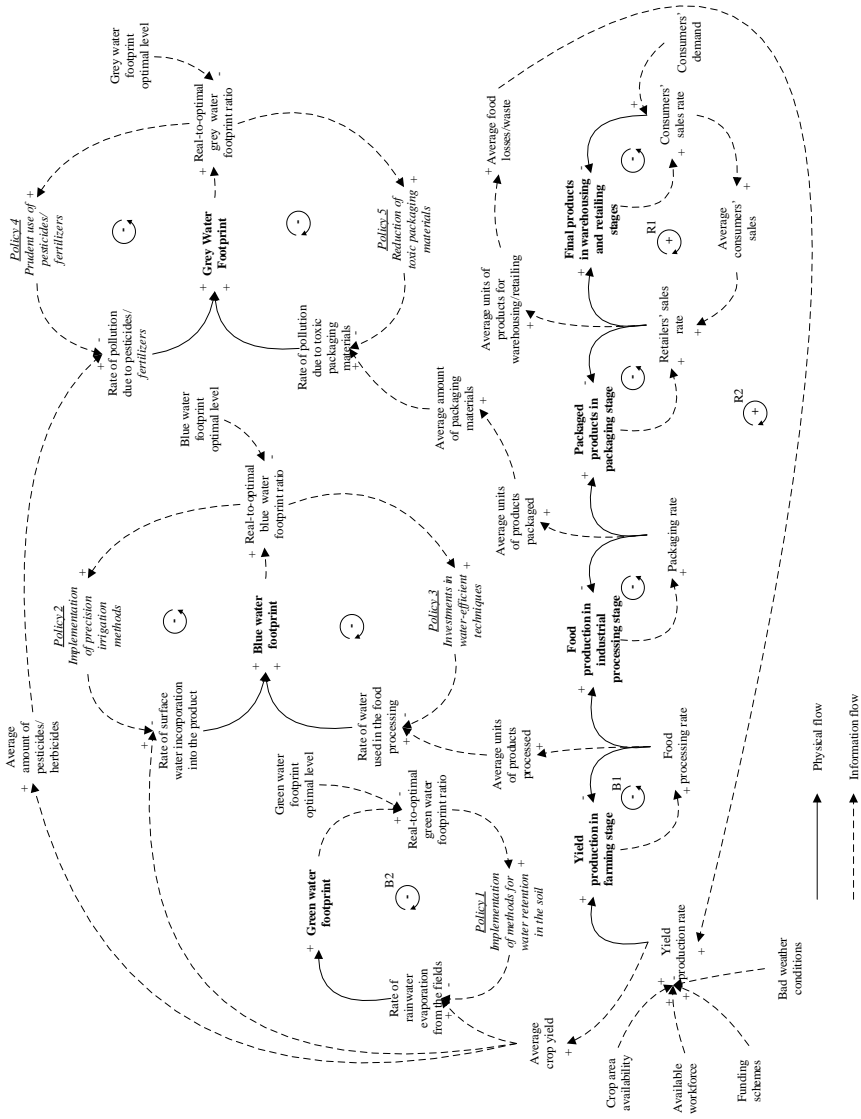
Fig. 3. SD model for WF management in AFSCs.

4   Discussion and Conclusions

The assessment and management of WF is becoming a key strategic issue for the contemporary businesses. Particularly, in the agrifood sector, the evaluation of the corporate WF has to be applied to all supply chain stages, namely from farming and industrial processing to packaging and distribution of the final products. At the same time, governments need to recognize the value of the domestic freshwater resources and implement sustainable supply chain management strategies and policies in order to ensure the production of sustainable and secure water-intensive agrifood commodities. At the same time, regulations and policies should assist in limiting the total WF due to regional problems of overexploitation and deterioration of freshwater resources.

In this context, our research reveals that the environmental impact of WF in the agrifood and industrial production sector varies significantly among EU-15 countries. The provided data along with the developed SD conceptual framework can contribute as a basis for policy-making concerning the sustainable management of freshwater supply across AFSCs in the European region. Particularly, the proposed WF management model is a first-effort strategic approach that can be used for supporting enterprises of the agrifood sector in limiting the green, blue and grey WF of their products through deploying sustainable policies in all supply chain stages.

Concerning our future research, this is going to be driven by real-world companies and supply chains towards the development of more robust SD models for the proper management of the WF across an AFSC. In addition, the EU-15 national WFs should be mathematically correlated with factors, such as freshwater availability, available crop land area, gross domestic product, as well as potential investments in water-efficient infrastructure in agriculture and industry. We expect that our work will assist in: (i) uncovering the critical factors affecting WFs that have not been examined systematically in the past, and (ii) providing significant managerial insights and policy-making recommendations for governments and business executives.

Acknowledgement

References

Aivazidou, E., Tsolakis N. and Iakovou, E. (2014) *Water Footprint Management for Agrifood Supply Chains: A Critical Taxonomy. Proceedings of the 3rd Pan-Hellenic Student Conference of the Hellenic Operational Research Society (HELORS)*. Piraeus, Greece.

Andersson, K., Ohlsson, T. and Olsson, P. (1998) Screening life cycle assessment (LCA) of tomato ketchup: A case study. *Journal of Cleaner Production*, 6, p. 277-288.

Ahmad, S. and Prashar, D. (2010) Evaluating municipal water conservation policies using a dynamic simulation model. *Water Resource Management*, 24(13), p. 3371-3395.

Ahmad, S. and Simonovic, S.P. (2000) System dynamics modeling of reservoir operations for flood management. *Journal of Computing in Civil Engineering*, 14(3), p. 190-198.

Ahmad, S. and Simonovic, S.P. (2006) An intelligent decision support system for management of floods. *Water Resource Management*, 20(3), p. 391-410.

Aviso, K.B., Tan, R.R., Culaba, A.B. and Cruz Jr., J.B (2011) Fuzzy input-output model for optimizing eco-industrial supply chains under water footprint constraints. *Journal of Cleaner Production*, 19(2-3), p. 187-196.

Boulay A., Hoekstra A.Y. and Vionnet S. (2013) Complementarities of Water-Focused Life Cycle Assessment and Water Footprint Assessment. *Environmental Science and Technology*, 47, p. 11926-11927.

Bourlakis, M., Maglaras, G., Gallear, D. and Fotopoulos, C. (2014) Examining sustainability performance in the supply chain: The case of the Greek dairy sector. *Industrial Marketing Management*, 43(1), p. 56-66.

Carter, C.R. and Dresner, M. (2001) Purchasing's role in environmental management: Cross-functional development of grounded theory. *Journal of Supply Chain Management*, 37(3), p. 12-27.

Ercin, A.E, Aldaya, M.M. and Hoekstra, A.Y. (2011) Corporate Water Footprint accounting and impact assessment: the case of the Water Footprint of a sugar-containing carbonated beverage. *Water Resources Management*, 25, p. 721-741.

Ercin, A.E, Aldaya, M.M. and Hoekstra, A.Y. (2012) The water footprint of soy milk and soy burger and equivalent animal products. *Ecological Indicators*, 18, p. 392-402.

Forrester, J.W. (1961) *Industrial Dynamics*. New York, USA: Productivity Press.

Forrester, J.W. (1969) *Urban Dynamics*. Cambridge, USA: MIT Press.

Gastelum, J.R., Valdes, J.B. and Stewart, S. (2010) A system dynamics model to evaluate temporary water transfers in the Mexican Conchos basin. *Water Resource Management*, 24(7), p. 1285-1311.

Gerbens-Leenes, P.W., Mekonnen, M.M. and Hoekstra, A.Y. (2013) The water footprint of poultry, pork and beef: A comparative study in different countries and production systems. *Water Resources and Industry*, 1-2, p. 25-36.

Herath, I., Green, S., Singh, R., Horne, D., van der Zijpp, S. and Clothier, B. (2013) Water footprinting of agricultural products: A hydrological assessment for the water footprint of New Zealand's wines. *Journal of Cleaner Production*, 41, p. 232-243.

Hoekstra, A.Y. (2003) *Virtual water trade: Proceedings of the International Expert Meeting on Virtual Water* Trade*, Value of Water Research Report Series No. 12*. Delft, The Netherlands: UNESCO-IHE.

Hoekstra, A.Y. (2008) *Water Neutral: Reducing and Offsetting the Impacts of Water Footprints, Value of Water Research Report Series No. 28*. Delft, the Netherlands: UNESCO-IHE.

Hoekstra, A.Y., Chapagain, A.K., Aldaya, M.M. and Mekonnen, M.M. (2011) *The Water Footprint Assessment Manual: Setting the Global Standard*. London, UK: Earthscan Ltd..

Hoekstra, A.Y., Gerbens-Leenes, P.W. and van den Meer, T.H. (2009) Reply to Pfister and Hellweg: Water footprint accounting, impact assessment, and life-cycle assessment. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 106(40), p. E114.

Hoekstra, A.Y. and Hung, P.Q. (2002) *Virtual water trade: A quantification of virtual water flows between nations in relation to international crop trade, Value of Water Research Report Series No. 11*. Delft, the Netherlands: UNESCO-IHE.

Hoekstra, A.Y. and Mekonnen, M.M. (2012) The water footprint of humanity. *Proceedings of the National Academy of Sciences*, 109(9), p. 3232-3237.

Iakovou, E., Vlachos, D., Achillas, C. and Anastasiadis, D. (2012) *A Methodological Framework for the Design of Green Supply Chains for the Agrifood Sector. Proceedings of the 2nd International Conference on Supply Chains (2nd Olympus ICSC)*. Katerini, Greece.

Khan S., Yufeng, L. and Ahmad, A. (2009) Analysing complex behaviour of hydrological systems through a system dynamics approach. *Environmental Modelling and Software*, 24(12), p. 1363-1372.

Klemeš, J.J., Varbanov, P.S. and Lam, H.L. (2009) Water footprint, water recycling and food-industry supply chains. In: K. Waldron (eds) *Handbook of Waste Management and Co-Product Recovery in Food Processing*. Cambridge, UK: Woodhead Publishing Series in Food Science, Technology and Nutrition, Woodhead Publishing Ltd., p. 134-168.

Liu, J., Yang, H. and Savenije, H.H.G. (2008) China's move to higher-meat diets hits water security. *Nature*, 454(7203), p. 397.

Manzardo, A., Ren, J., Piantella, A., Mazzi, A., Fedele, A. and Scipioni, A. (2014) Integration of water footprint accounting and costs for optimal chemical pulp supply mix in paper industry. *Journal of Cleaner Production*, 72, p. 167-173.

McKinsey (2009) *The global corporate water footprint: Risks, opportunities, and management options*. McKinsey & Company Inc.: CCSI Water & Adaptation Service Line.

Pegram, G., Orr, S. and Williams, C. (2009) *Investigating Shared Risk in Water: Corporate Engagement with the Public Policy Process*. Godalming, UK: WWF.

Qaiser, K., Ahmad, S., Johnson, W. and Batista, J. (2011) Evaluating the impact of water conservation on fate of outdoor water use: a study in an arid region. *Journal of Environmental Management*, 92(8), p. 2061-2068.

Roberts, E.B. (1978) *Managerial Applications of System Dynamics*. Cambridge, USA: Pegasus Communications Inc..

Rockstrom, J., Steffen, W., Noone, K., Persson, Å., Chapin, S., Lambin, E., Lenton, T., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., de Wit, C., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Synder, P., Costanza, R., Svendin, U., Falkenmark, M., Karlberg, L., Corell, R., Fabry, V., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P. and Foley, J. (2009) A safe operating space for humanity. *Nature*, 461, p. 472-475.

Rondinelli, D.A. and Berry, M.A. (2000) Environmental citizenship in multinational corporations: Social responsibility and sustainable development. *European Management Journal*, 18(1), p. 70-84.

Ruini, L., Marino, M., Pignatelli, S., Laio, F. and Ridolfi, L. (2013) Water footprint of a large-sized food company: The case of Barilla pasta production. *Water Resources and Industry*, 1-2, p. 7-24.

Simonovic, S.P. and Ahmad, S. (2005) Computer-based model for flood evacuation emergency planning. *Natural Hazards*, 34, p. 25-51.

Sterman, J. (2000) *Business Dynamics: Systems Thinking and Modelling for a Complex World*. Boston, USA: McGraw-Hill Higher Education.

Tidwell, V.C., Passell, H.D., Conrad, S.H. and Thomas, R.P. (2004) System dynamics modeling for community-based water planning: application to the Middle Rio Grande. *Aquatic Sciences-Research Across Boundaries*, 66(4), p. 357-372.

Tsolakis, N.K., Keramydas, C.A., Toka, A.K., Aidonis, D.A. and Iakovou, E.T. (2014) Agrifood supply chain management: A comprehensive hierarchical decision-making framework and a critical taxonomy. *Biosystems Engineering*, 120, p. 47-64.

UNESCO (2006) *Water a Shared Responsibility: the United Nations World Water Development Report 2*. Paris, France: UNESCO Publishing and New York, USA: Berghahn Books.

WBCSD (2006) *Business in the World of Water. WBCSD Water Scenarios to 2025*. Geneva, Switzerland: World Business Council for Sustainable Development.

Winz, I., Brierley, G. and Trowsdale, S. (2009) The use of system dynamics simulation in water resources management. *Water Resource Management*, 23(7), p. 1301-1323.

Yu, Z. and Qingshan, Y. (2014) Decoupling agricultural water consumption and environmental impact from crop production based on the water footprint method: A case study for the Heilongjiang land reclamation area, China. *Ecological Indicators*, 43, p. 29-35.

Zhang, X., Feng, H., Mao, X., Zheng, M. and Wang, R. (2009) *A dynamic water resources management approach in Beijing using system dynamics model*. Proceedings of the *International Conference on Management and Service Science (MASS 2009)*. Wuhan, China.

Comprehending HRM policies and practices in multinationals within the hospitality sector: country of origin and country of domicile effects

Giovanni Oscar Serafini[1]

[1]University of Sheffield Management School
City College Business Administration & Economics Dept.
South East European Research Centre,
giserafini@seerc.org

Abstract. This study focuses on human resources management (HRM) practices and policies applied in hotel multi-national enterprises (MNEs), which belong to a thriving and labor-intensive business area within services industry. By aiming to assess the determinants of human resource management application in different national contexts within a global luxury hotel chain, the research explores the relationship between corporate guidelines and application of HRM policies and practices to different national contexts, thus contributing to the debate on the nature and application of HRM in the hotel industry. Important findings could occur in relation with the degree of employer-employee interdependence as well as the amount of delegation to employees. By elaborating on a multiple case study methodology, we foresee that a hospitality MNE is able to ensure an overall uniform application of HRM policies and practices regardless of locale through a "geocentric" approach according to the famous model by Perlmutter.

Keywords: *home/host country effects; role of culture; role of institutions; luxury hotel industry.*

1 Introduction

This is a study aiming at analysing the application of Human Resource Management (HRM) policies in multi-national enterprises (MNEs) in the hospitality industry, while centring on the case of a US hotel chain. In particular, this research work is rooted in the literature on comparative capitalism, and it seeks to shed new light on the nature of institutional effects on firms that cross national boundaries. The focus is on the hotel industry, a rather labour intensive sector, but one that has been relatively neglected in the comparative study of MNE application of HRM policies and practices particularly because it legs behind in its implementation as compared to other industrial sectors (Lee-Ross and Pryce, 2010; Baum, 2007; Higgins-Desbiolles, 2006; Marchante et al., 2006). Besides, the literature on comparative capitalism has

tended to concentrate on the mature markets: this study brings to bear a new body of evidence from emerging markets as well.

Much of the literature on international HRM focuses on comparisons of a limited cross section of countries. This comparative case study research concerns a single multinational firm spanning over both developing and mature markets. Since it is recognized that there is much heterogeneity in the hospitality industry, focussing on a premium chain entails limitations. At the same time, this research allows a very much closer and direct comparison of variations in HRM policy according to locale than would otherwise be the case, within the same organization structure. In other words, it is possible to examine differences between settings within a single and mature organizational context. More broadly, hospitality services are an essential element of a tourism product related to a destination together with transportation, events, attractions, restaurants and bars. These, in turn, belong to a wider and complex tourism system (Riley, 2002) referred also as tourism market (Cooper, 2006, p. 183) that interacts with the customers through transactions and impacts. Consequently, as part of tourism markets, the hospitality industry is directly affected by the sheer seasonality of demand of the tourism product: this demand fluctuation often triggered by factors outside the suppliers' control affects the management of human resources within an industry which is particularly labour intensive owing to the nature of the service processes (Reisinger, 2001). The resulting hospitality-related employment opportunities create jobs and the income generated has a substantial economic consequence through a "multiplier" effect, especially if a destination relies only on tourism as a main economic activity (Pender, 2005, p. 8). Thus, by virtue of its link with travel and tourism, the hospitality industry has attained a global profile with an increasing importance in many national economies, thus placing local businesses in direct competition with large multinational hotel chains (Endo, 2006; Pine and Qi, 2004; Mathews, 2000).

Indeed, the hospitality industry represents a main indicator of future trends in terms of business cycle as it follows the development of markets worldwide: as an example, Asia, with special reference to China and India, the Pacific and Eastern Europe are forecast to become top tourism destinations by 2020 (Johnson and Vanetti, 2005; WTO, 2001), with the hospitality industry growing in those regions at a noteworthy rapid pace. It is estimated that over the next ten years (WTTC, 2011), the global travel and tourism economy will increase by 4.3% per year, leading its share of the global economy to just above 10%. In terms of jobs creation, the WTTC estimates that 66 million jobs will be added by 2020. While in 2010 the share of total workforce employed by the travel and tourism economy was 8.1%, this figure is expected to rise by 2020 to 9.2% of total employment. Even by considering the negative effects upon tourism development caused by the 9/11 terrorist attacks, flu disease breakouts, natural calamities, the WTO (2001) estimates that the 1 billion tourists' figure for 2010 would increase by 50% within just the next 10 years. Indeed, 2010 demonstrated a strong recovery after the drop of 4% recorded in 2009 caused by the global economic crisis. In real terms, international tourism arrivals in 2010

reached 935 million which represented an increase by almost 7% on the previous year (WTO, 2011).

Although recently more has been written about HRM in the hospitality industry (Davidson et al., 2011), nonetheless this has often occurred by utilizing mainstream HR research frameworks and models (Lucas and Deery, 2004). Further, owing to the vast, complex and multifaceted nature of the hospitality industry, Brotherton (1999) highlights the major challenge faced by academic researchers to appropriately determine the confines and scope of their field of research. Indeed, Guerrier and Deery (1998) note that most recent research on services industry relate to banking, airline and retail sectors and much less specifically hotels. Consequently, the potential contribution of this PhD lies in its effort to fill in a research void at two distinct levels. First, it is explicitly dedicated to the five-star global hotel industry, thus answers to the two-pronged need of more research in hotel organizations while ensuring clarity of scope within neat research boundaries. Second, following the observation by Drumm (1994) that HRM can be described as "some islands of theory in a large sea of untested hypotheses" (p. 46), this PhD study is expected to push further the frontiers of knowledge. In particular, this focussed single-company research aims to answer specific research questions (Redman and Mathews, 1998) thus offering a much-needed finer understanding of developments in international HRM (Geringer et al., 2002) applied in the hospitality industry. Based on the multiple case study methodology, we contend that luxury hospitality MNEs may ensure a uniform application of HRM policies and practices regardless of locale through a "geocentric" approach according to the Perlmutter model (1969).

2   Literature review

This research work aims to explore and understand commonality and differences in HRM practice within a luxury hotel firm straddling national boundaries. Thus, following an initial overview of the main challenges related to the implementation of HRM by MNEs, the literature review illustrates HRM occurring in the hospitality industry, before moving on to analyse the contemporary trends of thought related to elements affecting HRM in MNEs more broadly. In particular, key current debates reviewed encompass approaches to managing MNE subsidiaries (Holtbrügge and Mohr, 2011), cultural assumptions (Sartorius et al., 2011), institutional theoretical accounts (Brookes et al., 2011) as well as convergence and divergence explanations in relation to management systems within a globalized world (Paik et al., 2011).

2.1   Home and host country effects

MNEs are faced with a key dual challenge involving opposite forces that affect HRM policies and practices implementation: on the one hand, centralization and global

integration pushing for standardization, while, on the other, decentralization and local responsiveness pulling towards localization (Rosenzweig, 2006). From the previous sections, it derives that adaptation to local context is an imperative if a transnational firm is to implement a coherent strategy throughout its subsidiaries. The critical issue is, however, to what extent must a MNE adapt its country-of-origin practices to meet its host-country needs while it is demonstrated that home-country influences exert strong pressures on management systems, the culture of the organization as well as the application of human resource management processes (Lau and Ngo, 1996; Rosenzweig and Nohria, 1994). In particular, Taylor et al. (1996) indicate that MNEs, by drawing on their home-country competitive advantage, tend to implement homogeneous HRM systems in their subsidiaries overseas. The researchers show that this attempt is especially successful when the gap between the parent and host countries in terms of cultural and legal distance is limited.

According to Cyr and Schneider (1996), international joint ventures (IJVs) in transition economies of the former Eastern Bloc represent an interesting case in point whereby successful performance is the result of effective co-operation between locals and foreign nationals. As key prerequisites, the researchers indicate that "trust and mutual respect" (p. 219) are instrumental for the attainment of an enduring employee commitment, which, in turn enables the organization to learn and grow while relying on locally sound communication, training, staffing, appraisal and reward systems. Furthermore, in their recent research of Eastern European transition economies, Jindra et al. (2009) highlight that beneficial technological diffusion and spillovers towards the development of host-country economies are a direct consequence of a MNE subsidiary role and its technological competence. Again in this specific geographic area, which represents an ideal laboratory to test HRM implementation due to its recent Communist past, Lewis (2008) indicates that one of the most precious contribution of MNEs to their subsidiaries and society at large are the training and developmental opportunities offered to the local workforce while instilling the meanings of productivity and quality. As Kostova (1999) argues, the transfer of organizational practices within a MNE is effective provided that the social, organizational, and relational dimensions are fulfilled.

Equally important, Ferner and Edwards (1995) indicate that MNEs should also be considered as structures of power whereby operational outcomes are the effect of a dynamic interplay among resource possession and exchange, formal authority and vested interests. Within this viewpoint, the transfer of best practices is greatly facilitated by coercive comparisons which allow for the exertion of pressures on local management and workforce to align accordingly (Sisson et al., 2003). Nonetheless, Martin and Beaumont (1998) contend that MNEs should exert a viable pressure on their subsidiaries following the example of a model company such as ABB. More in detail, they remark that internal benchmarking should, on the one hand, consider the local context in terms of culture and institutions and, on the other hand, the capability and motives of local management to apply best practices.

From a different angle and in contrast with Minbaeva et al. (2003), Szulanski (1996) indicates that the diffusion of best practice within an organization has to be fostered by eliminating "internal stickiness" (p. 29) caused by elements beyond purely motivational aspects. Indeed, his research leads to the conclusion that barriers to knowledge transfer are caused, firstly, by the inability of the workforce to identify, appreciate and implement new knowledge; then, lack of depth of knowledge; and, last, by ineffective communication and poor interpersonal relationships. As Edwards (1998) contends, if there is prevalence of environmental elements favouring synergies and appropriate organizational "maturity" conditions in a global MNE, there may be cases of "reverse diffusion" (p. 696) according to which practices are even transferred from the international subsidiaries to other subsidiaries and domestic plants of the organization.

2.2   Institutions and home country effects

Ferner and Varul (2000) argue that HRM is critically affected by the country of origin model of personnel management, which is conditioned by factors within the home national business system. Their research outcome is the result of an analysis of the highly structured German institutional framework determining a considerably reactive and administration-centred personnel management approach pressing against the strategic adaptation of a wider HRM implementation along the lines of the 'Anglo-Saxon' model.

Parent country (i.e., home country or country-of-origin) effects represent a category of elements exerting pressures on the way multinationals operate, their organizational culture and, ultimately, their approach to HRM. Thus, MNEs from different home countries apply HRM in different ways for the very reason that, according to Ferner (1997):

> "even the most global of companies remain deeply rooted in the national business systems of their country of origin (…). Even where the home base does not account for the bulk of sales, operations and employment, the home nation is almost always the primary locus of ownership and control" (p. 19).

As illustrated by Chang et al. (2007) the pressures exerted onto a MNE in order to be aligned with home country institutional environments represent "push forces" (p. 405) and are the result of the MNE degree of embeddedness in its business system (Ferner, 1997). The home country effect can take many forms as demonstrated in numerous studies comparing practices in MNEs from different countries. In particular, nationality of ownership is a critical determinant of MNE business code of conduct and management practices following the comparison of US MNEs as opposed to Japanese and EU MNEs, for instance. While the former tend to be centralized and standardized in the management of HR, Japanese MNEs are more

keen to adapt to the local settings (Ferner, 1997). Also, the fact that MNEs mostly gravitate around their home country (Hejazi, 2007) is demonstrated in their concentration of assets, sales generated as well as the marked tendency to fill key positions at the executive level with parent country nationals (Edwards et al., 2007). The latter, in fact, plays a key role in the transfer process of knowledge and management practices as well as the approach to industrial relations by head office towards subsidiaries (Edwards et al., 2007).

Consequently, transfers for organizational development become a prerogative of head office nationals (Ondrack, 1985), a persistent ethnocentric focus by MNEs that Mayrhofer and Brewster (1996) vividly reinforce through comparative survey evidence. Indeed, there are some organizations which are very successful by adopting only parent country policies as happens, for instance, with US oil and gas firms as well Chinese construction MNEs whose operations are almost entirely staffed by expatriates (Corkin et al., 2008; McSherry, 2006; Wood, 2004). While in the former home-country expatriates are vital to ensure the skilful execution of highly technologically sophisticated operations, the latter are extremely successful as compared to other competitors owing to their low-cost but solidly trained and committed Chinese expatriate labor force. This organizational reality illustrates that ethnocentricity is one of the available avenues towards attaining convergence and internal consistency on a global scale (Lau and Ngo, 2001). Indeed, research highlights that MNEs are affected by parent country pressures towards a dominant global model on several key fronts, namely management system, corporate culture and HRM practices (Schuler and Rogovsky, 1998; Taylor et al., 1996; Rosenzweig and Nohria, 1994).

In the final analysis, it is critical to recognize the link parent country effects have with the issue of institutions as previously reviewed. In fact, as firms are located in the centre of a composite interaction web, their effective relationship with "producer groups, employees and other firms" (Hall and Thelen, 2009, p. 8) is directly affected by the support of institutions from a political economy standpoint. As a result and in proportion to its strength, also the parent country institutional framework influences the approach businesses have to their international operations.

2.3   Organizational isomorphism

The implementation of HRM by a MNE is not a straightforward process as it is conditioned by the modus operandi of subsidiaries and their dynamic interplay vis-à-vis the organization head office. In fact, Rosenzweig and Nohria (1994) in studying human resource management practices in 249 U.S. affiliates of foreign-based MNEs from Canada, Japan and Europe demonstrate that, broadly, affiliate HRM practices are strongly aligned with local practices while, however, differing in particular practices. The extent of "local isomorphism" (p. 241) as the authors put it, is

determined by the founding method, the reliance on local inputs, the ratio of expatriates within the workforce of the subsidiary, the closeness of communication with head office and the need for internal organizational consistency. According to DiMaggio and Powell (1983, p. 150) there are three main forms of isomorphic outcomes, each determined by particular processes: firstly, "coercive isomorphism" relates to the pressures onto an organization as occurs through the imposition of law, for instance; secondly, "mimetic isomorphism" indicates efforts of an organization to copy successful practices adopted by other organizations; and, last, "normative isomorphism" describes the adjustment of an organization to what is regarded as appropriate in a particular environment (Kostova and Roth, 2002; Haveman, 1993).

According to Jackson and Schuler (1995), local environment and culture have a key role in the application of HRM systems, policies and practices. On this same line, a study by Beechler and Yang (1994) on Japanese subsidiaries operating in the USA shows that local circumstances decisively influence a MNE's ability to transfer parent-country practices abroad. As a result of particular local constraints, businesses are mostly unable to accurately replicate home countries' practices in their international subsidiaries and thus opt for a locally isomorphic approach (Rosenzweig and Singh, 1991). Indeed, even between similar national business systems such as the American and the British, the transfer of HR practices from the US parent company to the UK subsidiary is impeded by local stakeholders who prevent a MNE to fully implement a corporate diversity policy, as shown in a case analysis by Ferner et al. (2005).

2.4 HRM in the hospitality industry

The hospitality industry, which is intertwined with the wider tourism industry, is one of the largest and most dynamic industrial sectors worldwide. In particular, the hospitality industry is identified with all businesses providing, according to Brotherton (1999):

> "a contemporaneous human exchange, which is voluntarily entered into, and designed to enhance the mutual wellbeing of the parties concerned through the provision of accommodation and food or drink." (p. 168)

The global expansion of tourism and of the hospitality industry within it is accompanied by two specific trends: on the one hand, consumers expect more quality and variety in the services and products they purchase while, on the other hand, competition among firms is heated up both nationally and internationally. Consequently, superiority in this industrial sector is attained by businesses which are able to surpass competition in terms of service quality, customer satisfaction and business performance (Augustyn and Ho, 1998). Thus, a key pre-requisite is the implementation of sound HRM practices owing to the fact that hospitality critically depends on successful human interactions (Davidson et al., 2011).

In order to duly illustrate the industrial sector under consideration, there is the need to identify the different business typologies. In fact, hospitality organizations share a most common services framework whereby accommodation is compounded with a number of supporting facilities such as specialty restaurants, shopping arcade, wellness centre, which can escalate to a very sophisticated range of services depending on the competition and the market addressed (King, 1995). For the purposes of this PhD, emphasis is given to MNEs operating in the luxury segment of the industry, otherwise referred to as "international hotel chains" which offer product-branded services on a standardized global basis (Whitla et al., 2007, p. 778).

2.5   HRM in the luxury hospitality industry MNEs

Same as occurs in any other industrial sector, hospitality firms do have the strategic option to expand their business operations internationally, or even globally. Contractor and Kundu (1998b) found that the entry mode into new markets is determined by country-, environment- and firm-specific variables. According to the researchers, 65.5% of hotel international operations are contracted through non-equity agreements, such as franchise agreements, management contracts and strategic alliances. Consequently, these arrangements enable hotel organizations with successful and unique service brands to expand globally. Further, a most recent study by Graf (2009) clearly indicates that stock markets react abnormally whenever the entry mode of the hotel chain fits with the specificities of the host country. In particular, the author found that investors react very favourably to new management contracts in developing countries and new franchise agreements in developed countries, also because of the anticipated benefits resulting from the hotel firms' strategic orientation, contractual control and formality of business operations (Yan et al., 2007). Notwithstanding, the successful implementation of HRM policies and practices in overseas subsidiaries depends on the sound interaction between owning entity and the hotel management company which either runs the business on their behalf through a management contract or allows the hotel to operate under a franchising agreement (Gannon et al., 2010).

For luxury hotel chains, emphasis on quality throughout is of utmost importance and, therefore, HRM plays a strategic role in strengthening this vision within the organization (Maxwell et al., 2004). In 1999 Hilton, for instance, introduced a worldwide customer service quality initiative named 'Equilibrium' which prompted HRM to develop a strategy branded 'Esprit' that encompassed, beyond a successful staffing function, policies and practices targeting specifically the areas of assessment, compensation, benefits, recognition, career tracking and discipline (Maxwell and Lyle, 2002). Other hotel chains, have put HRM at centre stage with regards to their business strategy along with marketing and operations as in the case of Accor (Aung, 2000), which, even, has implemented remarkable talent management initiatives through the dynamic involvement and support of the hotel General Manager themselves (Yeung, 2006).

Nonetheless, despite the global expansion of luxury hotel chains and the subsequent attempt to globalize the HRM function, a closer examination of local contexts demonstrates the existence of unique challenges pertaining to the political, national and cultural settings, which requires the local hotel HR professional to adapt accordingly (Naama et al., 2008; Costa, 2004; Lu and Chiang, 2003). Indeed, Boxall (1995) argues that differences in workforce capability, labour productivity and employment systems affect the management of human resources which needs to be modified from country to country.

Furthermore, labor markets critically affect the quality of service professionals: to this end, governments do contribute to the establishment of MNEs through, among others, the education of local nationals in hospitality vocational professions (Baum and Szivas, 2008; Kusluvan and Karamustafa, 2001). Even in developed countries such as Korea, research has shown that Training Managers working in international chain hotels modify training programs in order to make them fit to the local culture (Lim and Wentling, 1998). Equally important, labour shortage is a ubiquitous challenge to be found even in populous China, however with its distinct peculiarities related to lack of qualified labour force and the unwillingness for university graduates to join the industry (Zhang and Wu, 2004). On the other hand, surveys by Magnini and Honeycutt (2003) and Shay and Baack (2004) demonstrate the high failure rate of expatriated hotel managers as well as discrepancies between the expatriated managers' self-perception of performance and subordinate-rated managerial effectiveness.

In conclusion, local specificities compel HRM to adjust to the 'think global, act local' perspective whereby central offices originate wide-scope HRM policies and procedures and then it falls upon the local hotel HR executive to decide and act the best way they see fit to the local context which, incidentally, they know better than anybody else in the hierarchy of the HRM function (Enz, 2009). An example of such an approach is offered by Zuehl and Sherwyn (2001) who contend that, after analyzing employment termination practices in a sample of countries, MNEs feature common employment termination policies presented as general guidelines to be considered, but then empower the local HR executive to decide accordingly.

2.6   Differences in HRM implementation: hospitality MNEs v. local hotels

Although rising, multinational organizations still represent a sheer minority of all businesses operating in the hospitality industry: in fact, local small and medium sized hotel enterprises (SMEs) prevail all the world over as documented by Baum (1999) according to whom 90% of hospitality businesses globally belong to the SME type. As a result of their intrinsic lack of professional and business skills, insufficiency of funds and unstable business performance, most SMEs are evidently disadvantaged when it comes to competing on human resources management grounds with larger organizations (Baum, 1999).

Thus, the hospitality industry is theatre to a most striking contradiction: despite the fact that theoretical propositions and empirical evidences clearly indicate that HRM effectiveness is key to service quality, customer satisfaction, business sustainability and profitability as well as competitive advantage, still the reality of many local hotel organizations in terms of workforce management practices and employment conditions often demonstrate the lack of sound HRM (Cho et al., 2006). In practice, the role of HRM within such organizations is given low priority compared to other business functions such as finance and sales and marketing (McEvoy, 1984; Rutherford and O'Fallon, 2007; Haynes and Fryer, 2000; Kelliher and Johnson, 1997). This leads to questioning the credibility of the fundamental assumption widely expressed by hotel management according to which employees are the most important asset and key to success in relation to service quality, consumer satisfaction as well as business performance (Losekoot et al., 2001; Maher, 1993).

The reasons behind this inconsistency are varied, and can be traced to the labour-intensive character of the industry with a prevalent focus on external labour markets related to easily replaceable low-paid marginal workforce (Davidson et al., 2006; Bohle et al., 2004; Robson et al., 1999). This, coupled with high costs and thin profit margins in a markedly seasonal business environment (Jolliffe and Farnsworth, 2003), causes the sheer divide between theory and reality. It is no surprise, therefore, that the hospitality industry SMEs have a diffused poor reputation, particularly among hospitality and tourism students (Baum, 2002; Kusluvan and Kusluvan, 2000; Choy, 1995; Wood, 1992) because most often they are not a source of permanent and stable employment, offer uncompetitive retribution for working in anti-social hours while featuring few opportunities for job enrichment and career advancement.

On the other hand, hospitality MNEs with an international or, even, global reach have made remarkable efforts to narrow the gap between words and deeds and upgrade the profile of the industry by drawing from the long-term benefit of sound HRM practices implementation (Bohdanowicz and Zientara, 2009; Rowley and Purcell, 2001). As Baum (2006) maintains, it is vital for hospitality organizations to develop their workforce so as to render them "more flexible and adaptive to constant change" (p. 133) at both operational and managerial levels. Thus, Ford and Heaton (2001) contend that, due to the considerable intangible aspect of the service experience (Lashley, 1998), the hospitality industry has managed to develop unique competencies in service provision that businesses in any other industrial sectors may consider worth implementing. This because while the hotel industry is particularly capital- and labor-intensive, its logistics and supply chain processes can be as sophisticated as in manufacturing businesses (Dimou et al., 2003) contrarily to other sectors of the services industry, such as marketing and consulting.

Exactly because the hospitality industry is focused on providing service excellence in an extremely competitive environment (Dubéa and Renaghanb, 1999), hospitality MNEs aim to build a solid culture of service so as to guide every employee behavior by having the customer in mind and ensure that their experience is managed successfully (Teare, 1995). In turn, this reflects on managerial skills of MNE hospitality professionals which differ substantially from those required in the

manufacturing industry. Particularly, hospitality service delivery systems demand managers to develop special competencies in areas such as organizing, staffing and commanding (Bowen and Ford, 2004).

Consequently, being customer satisfaction and service delivery consistency the main strategic challenges faced by any hospitality concern, staffing is a top HRM priority with the goal to hire employees featuring a strong service attitude (Ladhari, 2009). Nonetheless, taking the local UK hotel market as an example, a survey by Price (1994) highlights that, despite these recruitment efforts, still personnel practices are poor for local hospitality SMEs in the mid-nineties as they were in the mid-eighties as further confirmed by Kelliher and Johnson (1987) and, later, by McGunnigle and Jameson (2000). However, in the mid-nineties that same local market indicates that large and foreign-owned hotels are featuring a more advanced approach to personnel management (Kelliher and Johnson, 1997), even if, overall, as compared with other business sectors in Great Britain, the hospitality industry features a remarkable divergence in the application of HRM. This is particularly highlighted by the markedly "retaining control/cost control" approach (Lucas, 2002, p. 211) to management leading to a very hard HRM which challenges employee motivation.

Another element worth considering is that the hospitality industry usually features a low level of unionization, whereby sound HRM strategy formulation and implementation critically depends on the effectiveness of management (Tanke, 2001). For instance, in the UK Lucas (1996) reports a decline in union membership in the 1980s from 6% to 3%, while union membership among hospitality workers in North America totals 14% (Tanke, 2001). The minor role of unions in the hospitality industry is due (Piso, 1999) to, first, the fact that the overwhelming majority of hospitality businesses belong to the national SME category characterized by a paternalistic style of management (Rowden, 1995); second, the varied and fragmented nature of hospitality industry (Ingram, 1999; Harrington and Akehurst, 1996); then, the very workforce composition does not facilitate the emergence of unions because of dissimilar employee profiles (young students, women, transient workers, minorities) and job contracts (part-time contracts, internships, outsourcing and subcontracting practices) (Soltani and Wilkinson, 2010); and last, weak internal job markets (Deery and Jago, 2002) as well as considerably high turnover rates (ILO, 2001; Price, 1994).

In the final analysis, as the research by Wilton (2006) reveals, the efficient and effective application of formal and strategic approaches to HRM is directly proportional to the size of the hospitality organization. This observation further strengthens the initial position that hospitality MNEs, owing to their vast resources and scale, are at a net advantage compared to national hotel SMEs in relation with the systematic implementation of sensible and innovative HRM practices (McPhail and Fisher, 2008).

2.7   Peculiarities of employment in the hospitality industry

In the hospitality service provision, the relationship between customer and employee is asymmetrical (Shamir, 1980) because it is the role of the latter to offer and of the former to receive, provided they want to (Jones and Sasser, 1995). The exchange is based on the fact that the guest must pay for the services with the understanding that they would not harm or disrupt the hotel, other customers as well as the service providers (Radolovic, 2010).   There are three key aspects to the host-guest relationship: first, the service provider does not hold a superior position in terms of a particular knowledge vis-à-vis the receiver as would happen with a doctor, an engineer or an economist; second, often the host is economically more affluent and of higher status than the service provider, something which is evident in luxury hospitality establishments (Goeldner and Ritchie, 2006); third, especially in customer-contact departments, tipping is an important source of income to the employees, something that gives the customer control of their behavior to a certain degree (Azar, 2007).   These basic elements plus the existence of intense competition put the customer in a clear position of advantage and control that presses hospitality organizations to forge in all employees an attitude of total customer satisfaction (Heart, 1988).

The main challenge faced by a hospitality professional are the many roles they need to cover at once (Dev and Olsen, 1989): on the one hand, while they aim to serve the needs of the customer they have still to focus on the control of the transaction so as to do it effectively; on the other, the hospitality professional needs to perform according to the organization brand standards, policies and procedures while maintaining their personal emotional balance and self-esteem.

Consequently, the function of human resources in the actualization and delivery of successful hospitality service experiences should not be underestimated: through effective HRM implementation as well as the nurturing of a people-centred organizational culture focussing on the respect of both the external and internal customers (employees), a hotel organization must aim to be truly hospitable. However, the issue still at stake is whether theory is applied in practice (Wilton, 2006).

2.8   Labour intensity

Boddewyn et al. (1986) recognized the importance of service MNEs as distinct from those non-service which deal with product manufacturing on the grounds that "by the year 2000, more than half of the world's multinational enterprises will be in services" (p. 54). Indeed, about twenty years later, a report by UNCTAD (2004) proved such forecast right, indicating the ascending role of workforce in the delivery of MNEs. Nonetheless, Lowell (1999) vividly contends that:

"World-class talent has enormous scale effects as interaction costs fall because it can be leveraged across ever greater geography. It underlies most opportunities to specialize. It is the ultimate source of all intangible capital. It underpins all the other components of intangible capital necessary for global success. And, unfortunately, it is in short supply" (p. 182).

It is a fact that the hospitality industry critically relies on its workforce as the products and services it offers are pervaded by the human element (King, 1995). As a result, contrarily to other industries such as manufacturing and agriculture, hospitality businesses feature high labour intensity which impedes radical downsizing because of the irreplaceability of service personnel (Nankervis, 2000). Indeed, the employment relationship between the hotel and its workforce is complicated by the very characteristics of the hospitality service: intangibility, perishability, heterogeneity, variability, simultaneity of production and consumption as well as inseparability (Reisinger, 2001). Overall, experts maintain that the level of labour intensity is also affected by the extent of tourism development of a particular destination and its infrastructure as well as the type of business (Baum, 2006).

Consequently, even if at varying degrees, labour expenses by themselves represent a considerable share of the total production costs all over the world in an industry with high cyclicality and tight profit margins (Woodworth and Mandelbaum, 2010). This results in a compression of the remuneration of employees to levels below industry average and in the implementation of practices to contain, if not reduce, payroll cost. Thus, hotel companies often resort to outsourcing services in order to seek a more just-in-time labour provision (Gonzalez et al., 2011; Lamminmaki, 2011) together with the implementation of part-time and casual work contracts (Cairncross and Kelly, 2008).

2.9   Employee turnover challenge

The above observations suggest a most challenging reality which is represented by the instability of labour resulting in alarmingly high turnover rates especially at the lower ranks (Wood, 1992). In fact, some researchers have even commented that the hospitality industry features a "labour turnover culture" (Davidson et al., 2010, p.452). Indeed, an International Labour Organization report on the hotel, catering and tourism sector (ILO, 2001) shows that in 1997 the U.S. feature a 51.7% turnover rate for line employees compared to 30% in Asia and 42% in the UK. To illustrate how this compares with other industrial sectors, in the case of the UK, hospitality industry turnover is second only to retail trade with 43.5% and far ahead of construction with 25% (ILO, 2001). The ILO additionally highlights that on average 25% of the total workforce is on a part time contract in 15 EU countries, while the figure rises to about 50% for Northern countries such as the UK, Denmark and the Netherlands. As Iverson and Deery (1997) maintain, this overarching turnover culture is further exacerbated by sheer seasonal fluctuations of tourism demand as well as unforeseen and uncontrollable circumstances of social, economic and political nature.

Interestingly, as the hospitality industry offers employment ranging from the unskilled to the highly skilled, its workforce composition features a contrast between employees at the periphery against those at the core (Brien, 2010; Knox, 2010). This particular differentiation of workforce depending on their indispensability is reflected on their status within the organization as well as on their job satisfaction which, in turn, is mirrored in their relative turnover rates. In fact, the ILO (2001) statistics confirm that in 1997, turnover rates in the US lodging industry for supervisors and property managers are about a fifth of those related to line employees as noted before. Further, the research of Lai et al. (2008) indicates that peripheral hotel employees are often considered costs rather than resources. Thus, it can be inferred that the difference between a hotel professional and an employee is to be found in the fact that the former, as opposed to the latter, identify themselves in their profession and enjoy more stable working conditions while the latter is exposed to contingent labour arrangements owing to their disposability (Hjalager and Andersen, 2001).

Nonetheless, the hospitality industry (Hui and Hsin-Wei, 2009) has its own share of responsibility in leading turnover to such high levels because of the widespread poor working conditions and human resources practices. This includes unattractive retribution, absence of career planning, uncertainty of employment as well as long and unsocial working hours (Boella and Goss-Turner, 2005) particularly in local hotel businesses that may even reach Orwellian connotations in certain cases (Orwell, 1933). The result is that many qualified employees, after a short stint in an industry featuring mostly seasonal and part-time employment, seek better professional opportunities elsewhere. Thus, hotels are especially attractive employers to the "secondary labour market" (Price, 1994, p. 47), who are not distinctly dedicated to a specific industry such as students, housewives, school leavers and unqualified workers. These job-seekers search for a convenient source of income even on a part-time basis (Walsh, 1990) who, while having limited alternative occupational opportunities, appreciate working in contact with people (Curtis and Lucas, 2001).

As can be inferred from above, literature documentation (Deery and Shaw, 1999) indicates that employee turnover has both positive and negative consequences. On the positive side, it can be argued that it offers manning flexibility to serve effectively a fluctuating customer demand (Graham and Lennon, 2002). On the negative side, uncontrollably high turnover levels cause stress, damage employee morale, raise the workload to incumbent workforce due to systemic staff shortages and skills, impede the development and improvement of products and services, increase replacement costs, eliminate the benefits of training, with the end result of fostering a disengaged employee attitude leading to a drop in customer satisfaction and ultimately, to loss of revenue and profitability (Lee and Way, 2010).

2.10   Nature of hospitality jobs and related labour markets

Statistics reveal that an overwhelming share of jobs in the hospitality industry is unskilled or semi-skilled and that approximately 36% of jobs relate to managerial, supervisory and skilled (craft) positions (Riley, 1996). This reality has favoured society's widespread perception that jobs in hospitality are of low quality resulting in a downward pressure in terms of compensation and working conditions (Simons and Enz, 1995). Given the practical nature of the business, managers in the hospitality industry give emphasis on the operational skills and on-the-job training capability of the workforce, something that does not necessarily demand particular qualifications (ILO, 2001). Thus, such circumstances featuring both de-skilling and standardization in the name of a McDonaldization of processes (Lashley, 2008; Ritzer, 1998) facilitate the transferability of skills within the industry which justify also the high employee mobility within the industry (Deery and Jago, 2002).

At this point a clear distinction needs to be made between employment needs in developing as opposed to developed countries (Baum, 1996). In developing countries, hospitality industry is compelled to recruit at a considerably higher level than it would occur in developed countries owing to the shortage of local workforce who is educated and exposed to the outward culture of tourism and hospitality. The very need to fill the gap related to cultural and communication skills combined with the imperative to ensure world-class service standards, oblige industry operators in developing countries to raise the workforce profile by recruiting overseas. The result is that, according to Baum (1996), in many developing countries the ratio of semi-skilled or unskilled jobs is reduced to 15-25 % of the workforce in the hospitality industry.

By taking into consideration the high labour intensity of the hospitality industry, its considerable turnover rate as well as the critical share of unskilled and semi-skilled jobs over total workforce, it can be argued that the hospitality industry is generally characterized by a weak internal labour market. In fact, the particular nature of the business prevents the majority of establishments to foster a well-organized work environment with clear standards in terms of job description, performance evaluation, pay and promotion. According to research (Woods, 1999), employment in the hospitality industry is especially sought by women thus determining a feminization of the work whereby they would cover mainly mothering roles (housekeepers, waiting staff) or glamorous roles (front desk positions, sales and marketing associates). Nonetheless, developing countries may not feature a pronounced workforce feminization in their hospitality industry due to cultural reasons and women's position within the society according to which it is custom to have them work behind the scenes or in a back-of-the-house position (Burrell et al., 1997). Consequently, it can be argued that the status of employment in the hospitality industry in a country or region is affected by the general economic, social and cultural settings, the degree of

economic diversity as well as the level of dependence on the tourism activity (Matias et al., 2009). In the final analysis, this study addresses literature gaps at multiple levels. Initially, it investigates the hospitality industry that, even if it represents an important economic sector through its FDI impact, is neglected by mainstream literature. Then, it involves a particularly labour-intensive industrial sector that presents much interest from the standpoint of HRM analysis. Last, and most importantly, this comparative in-depth case study further sheds light on the application of HRM in post-Soviet economies which still remains largely unexplored by scholars.

3 Methodology

The company under consideration is a global deluxe hotel chain that, for the purpose of safeguarding anonymity, is referred to as HotelCo. This offers a unique research opportunity to understand the influence exerted by the MNE centre towards the periphery in an industry where MNEs are more spatially bound than many other employers. In fact, the peculiarity of international hotel chains is that they cannot, unlike manufacturing and some services such as call centres, move to the locations of lowest possible cost in a race-to-the-bottom strategic choice as previously explored (Contractor et al., 2010; Wright et al., 2005). Rather, they have to base their decisions primarily on consumer demand, resulting in hotel chains being present in some extremely high wage and heavily regulated contexts (Contrepois and Jefferys, 2010; Knox, 2010; McDowell et al., 2007). This makes hotels a very interesting case because they cannot simply exit markets to avoid regulation as a blanket policy since they need to be where customers are (albeit that they may choose to leave a specific market if the costs of labour, for example, are so great as to make for uncompetitiveness), but rather optimize their strategy according to the particular context and competition settings (Timo and Davidson, 2005).

The choice of a "collective case study" (Stake, 2000, p. 437) enables to identify similarities and dissimilarities across responses. The fact that research participants belong to the same MNE adds to the comparability of results among hotel subsidiaries in transitional periphery and advanced economies (Cooke, Wood, Psychogios, & Szamosi, 2011). Following Welch, Piekkari, Plakoyiannaki, and Paavilainen-Mantymaki (2011), case study methodology is instrumental to international business "given that as a field its raison d'être is to explain phenomena in diverse national, cultural and institutional contexts" (p. 756).

In the final analysis, using a case study organization that operates globally within the fast growing service sector it is believed it will enable the exploration of the relationship between corporate guidelines and application of HRM policies and practices to different national contexts, thus contributing to the debate on the nature and application of HRM in the hotel industry.

4 Empirical analysis

With the overall aim of this PhD research being the understanding of the application of HRM policies and practices by a US hotel MNE both in its home and foreign subsidiaries from the viewpoint of HR executives, research objectives are to:

1. identify, within the context of the luxury hospitality industry, the degree of autonomy overseas subsidiary HR executives have in setting their HR policies versus head office guidelines;
2. compare and contrast the HR practices encountered across overseas subsidiaries of the case study firm;
3. explore the wider theoretical implications of any homogeneity or diversity within the parameters of business systems theory, and the lessons for applied HR policy and practice.

The applied contribution to knowledge of this particular study can be understood by considering that it:

- highlights differences and similarities in the practice of people management in multinational hotel chains in different national contexts;
- offers a deeper understanding of the interdependence between MNE head office and HR executives and the level of delegation to the latter within a leading luxury hospitality MNE operating in a quintessential people business;
- illustrates common features which may constitute best practices;
- sheds further light on the constraints on innovation in HRM practice in different national settings, and explores ways in which these constraints may be overcome;
- contributes, at a theoretical level, to the emerging body of literature that applies comparative capitalism to the study of MNEs;
- describes the relative strength of country of origin versus country of domicile pressures in a service sector industry;
- provides further insights on the nature of institutional effects in emerging versus developed market settings.

The research is set to be implemented across a selection of HotelCo's overseas properties in the widely implemented "cross-sectional" fashion thus providing a "snapshot" at a particular point in time of specific occurrences and their connections (Saunders et al., 2007, p. 148), something that can be effectively attained via the use of a questionnaire. Consequently, a drawback of this choice is that it prevents from obtaining data over a given time period. This would have happened with a "longitudinal" approach allowing for a "causal inference" as is the case with studies related to change (Birdi et al., 2008, p. 476).

As the research is planned to occur at two sequential stages, initially data are collected and analysed quantitatively through the administration of an e-mailed survey subdivided in the following sections:

- autonomy in relation with HRM decisions
- communication and information flows
- employer-employee interdependence
- delegation to employees
- host-country specific factors determining the application of human resource management practices in multinational hotel chain subsidiaries:
  - selection practices
  - compensation practices
  - performance appraisal practices
- about expatriates
- workplace information and demographics

Next, additional data are gathered and examined qualitatively through individual telephone interviews following a semi-structured list of questions. Such a strategic approach reflects the "mixed method research" (Saunders et al,.2007, p. 145) applied to this intensive comparative case study research. More in detail, information is sought in order to assess the extent of formal corporate HRM strategy application by subsidiaries, drawing from the input via questionnaire administration to HR executives operating at overseas subsidiary level in seven properties located in the EAME (Europe-Africa-Middle East) division. The fact that participants to the survey are Human Resources executives provides a unique insight into the process directly from the management representatives in charge of implementing HRM policies and practices at hotel unit level. The wealth of information obtained allows for direct comparison across subsidiaries (Pudelko and Harzing, 2007).

5   Conclusion

This article overviews a doctoral research work in progress which is focussed on the field of HRM applied to the global luxury hotel industry. The review of literature indicates that business and context specificities including host country effects, the role of culture and institutions plus globalization pressures affect the application of HRM in practice across countries where a MNE operates.

Through a cross-national comparative analysis, the study enables the description of HRM implementation in seven cases related to subsidiaries of a luxury hotel MNE operating in Coordinated Market (CME), Liberal Market (LME) and Transitional Periphery Economies (TPE). The input from senior field HR professionals allowed the examination of HRM topics along the lines defined by Whitley's (1999) two core dimensions of work and employment relations. Following the institutional strand of

literature, a detailed cross-national comparative analysis indicates that MNE subsidiaries located in the two transitional periphery economies of Azerbaijan and Kyrgyzstan feature many similarities owing to the institutional settings that are still considerably influenced by their Soviet past, namely labour laws and education. Further, striking differences between TPE on one hand, and CME and LME on the other, pertain to employment flexibility and education level. The role of head office support as well as the engagement of expatriates and management towards a "culture that emphasizes the value of people" (Hinkin & Tracey, 2010, p. 160) emerged from the case study analysis as critical determinants in ensuring that policies and procedures are consistently applied regardless of location according to a geocentric approach (Yu, Byun, & Lee, 2014). Being the hotel group a global organization, the research is expected to offer useful insight on how HRM is occurring in different countries and contexts thus allowing the interpretation of results across locations and economies. Although this single industry analysis has its merits, it cannot be easily generalized owing to the small sample of respondents: thus, caution is advised in drawing broad conclusions.

References

Augustyn, M. and Ho, S. K. (1998) Service quality and tourism. *Journal of Travel Research,* 37 (1), p. 71-75.

Aung, M. (2000) The Accor multinational hotel chain in an emerging market: through the lens of the core competency concept. *The Service Industries Journal,* 20 (3), p. 43-60.

Azar, O. H. (2007) Why pay extra? Tipping and the importance of social norms and feelings in economic theory. *Journal of Socio-Economics,* 36 (2), p. 250-265.

Baum, T. (1996) Unskilled work and the hospitality industry: myth or reality? *International Journal of Hospitality Management,* 15 (3), p. 207-209.

Baum, T. (1999) Human resource management in tourism's small business sector: policy dimensions. *In*: Lee-Ross, D. (ed.) *HRM in tourism and hospitality: international perspectives on small to medium sized enterprises*. London, Cassell, p. 3-16.

Baum, T. (2002) Skills training in the hospitality sector: a review of issues. *Journal of Vocational Education & Training,* 54 (3), p. 343-363.

Baum, T. (2006) The future of work and employment in tourism. *In*: Buhalis, D. and Costa, C. (eds.) *Tourism management dynamics: trends, management and tools.* Oxford, Elsevier Butterworth-Heinemann, p. 130-136.

Baum, T. (2007) Human resources in tourism: Still waiting for change. *Tourism Management* 28 (6), p. 1383-1399.

Baum, T. and Szivas, E. (2008) HRD in tourism: a role for government? *Tourism Management,* 29 (4), p. 783-794.

Beechler, S. and Yang, J. Z. (1994) The transfer of Japanese-style management to American subsidiaries: contingencies, constraints, and competencies. *Journal of International Business Studies,* 25 (3), p. 467-491.

Birdi, K., et al. (2008) The impact of human resource and operational management practices on company productivity: a longitudinal study. *Personnel Psychology,* 61 (3), p. 467-501.

Boddewyn, J. J., et al. (1986) Service multinationals: conceptualization, measurement and theory. *Journal of International Business Studies,* 17 (3), p. 41-57.

Boella, M. J. and Goss-Turner, S. (2005) *Human resource management in the hospitality industry: an introductory guide.* Oxford, Elsevier Butterworh-Heinemann.

Bohdanowicz, P. and Zientara, P. (2009) Hotel companies' contribution to improving the quality of life of local communities and the well-being of their employees. *Tourism and Hospitality Research* 9(2), p. 147-158.

Bohle, P., et al. (2004) Working hours, work-life conflict and health in precarious and "permanent" employment. *Revista de Salude Publica,* 38 p. 19-25.

Bowen, J. and Ford, R. C. (2004) What experts say about managing hospitality service delivery systems. *International Journal of Contemporary Hospitality Management,* 16 (7), p. 394-401.

Boxall, P. (1995) Building the theory of comparative HRM. *Human Resource Management Journal,* 5 (5), p. 5-17.

Brien, A. (2010) Managing contingent labour in the hotel Industry by developing organisational social capital. *Journal of Human Resources in Hospitality & Tourism,* 9 (2), p. 163 - 178.

Brookes, M., et al. (2011) Measuring competing explanations of human resource management practices through the Cranet survey: cultural versus institutional explanations. *Human Resource Management Review,* 21 (1), p. 68-79.

Brotherton, B. (1999) Towards a definitive view of the nature of hospitality and hospitality management. *International Journal of Contemporary Hospitality Management,* 11 (4), p. 165-173.

Burrell, J., et al. (1997) Equal opportunities for women employees in the hospitality industry: a comparison between France, Italy, Spain and the UK. *International Journal of Hospitality Management,* 16 (2), p. 161-179.

Cairncross, G. and Kelly, S. (2008) Human resource development and 'casualisation' in hotels and resorts in eastern Australia: getting the best to the customer? *Journal of Management & Organization,* 14 (4), p. 367-384.

Chang, Y. Y., et al. (2007) HRM strategies and MNCs from emerging economies in the UK. *European Business Review,* 19 (5), p. 404-419.

Cho, S., et al. (2006) Measuring the impact of human resource management practices on hospitality firms' performances. *International Journal of Hospitality Management,* 25 (2), p. 262-277.

Choy, D. (1995) The quality of tourism employment. *Tourism Management,* 16 (2), p. 129-137.

Contractor, F. J., et al. (2010) Reconceptualizing the firm in a world of outsourcing and offshoring: the organizational and geographical relocation of high-value company functions. *Journal of Management Studies,* 47 (8), p. 1417-1433.

Contractor, F. J. and Kundu, S. K. (1998b) Modal choice in a world of alliances: analyzing organizational forms in the international hotel sector. *Journal of International Business Studies,* 29 (2), p. 325-357.

Contrepois, S. and Jefferys, S. (2010) European Works Councils in Central and Eastern Europe: varieties of institution building among French service sector multinationals. *Industrial Relations Journal,* 41 (6), p. 584-602.

Cooke, F. L., Wood, G., Psychogios, A. G., & Szamosi, L. T. (2011). HRM in emergent market economies: Evidence and implications from Europe. *Human Resource Management Journal, 21*(4), 368-378.

Cooper, C., et al. (2006) New and emerging markets. *In*: Buhalis, D. and Costa, C. (eds.) *Tourism business frontiers: consumers, products and industry*. Oxford, Elsevier Butterworth-Heinemann, p. 19-29.

Corkin, L., et al. (2008) *China's role in the development of Africa's infrastructure*. [online]. Centre for Chinese Studies, Stellenbosch University. Available from: http://www.sais-jhu.edu/bin/e/j/CorkinetalWorkingPaper.pdf [Accessed 25 / 02 / 2011].

Costa, J. (2004) The Portuguese tourist sector: key challenges for human resources management. *International Journal of Contemporary Hospitality Management,* 16 (7), p. 402-407.

Curtis, S. and Lucas, R. (2001) A coincidence of needs? Employers and full-time students. *Employee Relations,* 23 (1), p. 38-54.

Cyr, D. J. and Schneider, S. C. (1996) Implications for learning: Human Resource Management in East-West joint ventures. *Organization Studies,* 17 (2), p. 207-226.

Davidson, M., et al. (2006) Employment, flexibility and labour market practices of domestic and MNC chain luxury hotels in Australia: Where has accountability gone? *International Journal of Hospitality Management,* 25 (2), p. 193-210.

Davidson, M. C., et al. (2010) How much does labour turnover cost? A case study of Australian four- and five-star hotels. *International Journal of Contemporary Hospitality Management,* 22 (4), p. 451-466.

Davidson, M. C. G., et al. (2011) Hospitality HRM: past, present and the future. *International Journal of Contemporary Hospitality Management Volume,* 23 (4), p. 498-516.

Deery, M. and Jago, L. K. (2002) The core and the periphery: an examination of the flexible workforce model in the hotel industry. *International Journal of Hospitality Management,* 21 (4), p. 339-351.

Deery, M. A. and Shaw, R. N. (1999) An investigation of the relationship between employee turnover and organizational culture. *Journal of Hospitality & Tourism Research,* 23 (4), p. 387-400.

Dev, C. S. and Olsen, M. D. (1989) Applying role theory in developing a framework for the management of customer interactions in hospitality businesses. *International Journal of Hospitality Management,* 8 (1), p. 19-33.

DiMaggio, P. J. and Powell, W. W. (1983) The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review,* 48 (2), p. 147-160.

Dimou, I., et al. (2003) Corporate development of international hotel firms: a transaction cost economics and agency theory approach. *International Journal of Tourism Research,* 5 (2), p. 147-149.

Drumm, H. J. (1994) Theoretical and ethical foundations of human resource management: a German point of view. *Employee Relations,* 16 (4), p. pp.35-47.

Dubéa, L. and Renaghanb, L. M. (1999) A look at the industry's overall best-practice champions. *The Cornell Hotel and Restaurant Administration Quarterly,* 40 (6), p. 16-26.

Edwards, T. (1998) Multinationals, labour management and the process of reverse diffusion: A case study. *The International Journal of Human Resource Management,* 9 (4), p. 696 - 709.

Edwards, T., et al. (2007) Conceptual approaches to the transfer of employment practices in multinational companies: an integrated approach. *Human Resource Management Journal,* 17 (3), p. 201-217.

Endo, K. (2006) Foreign direct investment in tourism - flows and volumes. *Tourism Management,* 27 (4), p. 600-614.

Enz, C. A. (2009) Human Resource Management: a troubling issue for the global hotel industry *Cornell Hospitality Quarterly,* 50 (4), p. 578-583.

Ferner, A. (1997) Country of origin effects and HRM in multinational companies. *Human Resource Management Journal,* 7 (1), p. 19-37.

Ferner, A., et al. (2005) Institutional theory and the cross-national transfer of employment policy: the case of 'workforce diversity' in US multinationals. *Journal of International Business Studies,* 36 (3), p. 304-321.

Ferner, A. and Edwards, P. (1995) Power and the diffusion of organizational change within multinational enterprises. *European Journal of Industrial Relations,* 1 (2), p. 229-257.

Ferner, A. and Varul, M. Z. (2000) Internationalisation and the personnel function in German multinationals. *Human Resource Management Journal,* 10 (3), p. 79-96.

Ford, R. C. and Heaton, C. P. (2001) Lessons from hospitality that can serve anyone. *Organizational Dynamics,* 30 (1), p. 30-47.

Gannon, J., et al. (2010) The impact of hotel management contracting on IHRM practices: understanding the bricks and brains split. *International Journal of Contemporary Hospitality Management,* 22 (5), p. 638-658.

Geringer, J. M., et al. (2002) In search of "best practices" in international human resource management: research design and methodology. *Human Resource Management,* 41 (1), p. 5-30.

Goeldner, C. and Ritchie, B. (2006) *Tourism: principles, practices, philosophies.* New York, Wiley

Gonzalez, R., et al. (2011) What do we know about outsourcing in hotels? *The Service Industries Journal,* 31 (10), p. 1669 - 1682.

Graf, N. S. (2009) Stock market reactions to entry mode choices of multinational hotel firms. *International Journal of Hospitality Management,* 28 (2), p. 236-244.

Graham, M. and Lennon, J. (2002) The dilemma of operating a strategic approach to human resource management in the Scottish visitor attraction sector. *International Journal of Contemporary Hospitality Management,* 14 (5), p. 213-220.

Guerrier, Y. and Deery, M. (1998) Research in hospitality human resource management and organizational behaviour. *International Journal of Hospitality Management,* 17 (2), p. 145-160.

Hall, P. A. and Thelen, K. (2009) Institutional change in varieties of capitalism. *Socio-Economic Review,* 7 (1), p. 7-34.

Harrington, D. and Akehurst, G. (1996) Service quality and business performance in the UK hotel industry. *International Journal of Hospitality Management,* 15 (3), p. 283-298.

Haveman, H. A. (1993) Follow the leader: mimetic isomorphism and entry into new markets. *Administrative Science Quarterly,* 38 (4), p. 593-627.

Haynes, P. and Fryer, G. (2000) Human resources, service quality and performance: a case study. *International Journal of Contemporary Hospitality Management,* 12 (4), p. 240-248.

Heart, C. (1988) The power of unconditional service guarantees. *Harvard Business Review,* (7/8), p. 55-61.

Hejazi, W. (2007) Reconsidering the concentration of US MNE activity: Is it global, regional or national? *Management International Review,* 47 (1), p. 5-27.

Higgins-Desbiolles, F. (2006) More than an "industry": The forgotten power of tourism as a social force. *Tourism Management*, 27 (6), p. 1192-1208.

Hjalager, A.-M. and Andersen, S. (2001) Tourism employment: contingent work or professional career? *Employee Relations,* 23 (2), p. 115-129.

Hinkin, T. R., & Tracey, J. B. (2010). What makes it so great? An analysis of human resources practices among fortune's best companies to work for. *Cornell Hospitality Quarterly, 51*(2), 158-170.

Holtbrügge, D. and Mohr, A. (2011) Subsidiary interdependencies and international human resource management practices in German MNCs. *Management International Review,* 51 (1), p. 93-115.

Hui, O. Y. and Hsin-Wei, F. (2009) Contemporary human resource management issues and concerns in the hotel industry: identifying the priorities. *International Journal of Organizational Innovation,* 2 (1), p. 201-224.

ILO (2001) *Report for discussion at the tripartite meeting on human resources development, employment and globalization in the hotel, catering and tourism sector.* [online]. International Labour Office, Geneva. Available from: http://www.ilo.org/public/english/dialogue/sector/techmeet/tmhct01/tmhct-r.pdf [Accessed 26th November 2010].

Ingram, H. (1999) Hospitality: a framework for a millennial review. *International Journal of Contemporary Hospitality Management,* 11 (4), p. 140-148.

Iverson, R. D. and Deery, M. (1997) Turnover culture in the hospitality industry. *Human Resource Management Journal,* 7 (4), p. 71-82.

Jackson, S. E. and Schuler, R. S. (1995) Understanding human resource management in the context of organizations and their environments. *Annual Review of Psychology,* 46 (1), p. 237-264.

Jindra, B., et al. (2009) Subsidiary roles, vertical linkages and economic development: lessons from transition economies. *Journal of World Business,* 44 (2), p. 167-179.

Johnson, C. and Vanetti, M. (2005) Locational strategies of international hotel chains. *Annals of Tourism Research,* 32 (4), p. 1077-1099.

Jolliffe, L. and Farnsworth, R. (2003) Seasonality in tourism employment: human resource challenges. *International Journal of Contemporary Hospitality Management,* 15 (6), p. 312-316.

Jones, T. and Sasser, W. (1995) Why satisfied consumers defect. *Harvard Business Review,* 73 (6), p. 88-99.

Kelliher, C. and Johnson, K. (1987) Personnel management in hotels - some empirical observations. *International Journal of Hospitality Management,* 6 (2), p. 103-108.

Kelliher, C. and Johnson, K. (1997) Personnel management in hotels – an update: a move to human resource management. *Progress in Tourism and Hospitality Research,* 3 p. 321-331.

King, C. A. (1995) What is hospitality? *International Journal of Hospitality Management,* 14 (3-4), p. 219-234.

Knox, A. (2010) 'Lost in translation': an analysis of temporary work agency employment in hotels. *Work, Employment & Society,* 24 (3), p. 449-467.

Kostova, T. (1999) Transnational transfer of strategic organizational practices: a contextual perspective. *The Academy of Management Review,* 24 (2), p. 308-324.

Kostova, T. and Roth, K. (2002) Adoption of an organizational practice by subsidiaries of multinational corporations: institutional and relational effects. *The Academy of Management Journal,* 45 (1), p. 215-233.

Kusluvan, S. and Karamustafa, K. (2001) Multinational hotel development in developing countries: an exploratory analysis of critical policy issues. *International Journal of Tourism Research,* 3 (3), p. 179-197.

Kusluvan, S. and Kusluvan, Z. (2000) Perceptions and attitudes of undergraduate tourism students towards working in the tourism industry in Turkey. *Tourism Management,* 21 p. 251-269.

Ladhari, R. (2009) Service quality, emotional satisfaction, and behavioural intentions: a study in the hotel industry. *Managing Service Quality,* 19 (3), p. 308-331.

Lai, P.-C., et al. (2008) Distancing flexibility in the hotel industry: the role of employment agencies as labour suppliers. *The International Journal of Human Resource Management,* 19 (1), p. 132-152.

Lamminmaki, D. (2011) An examination of factors motivating hotel outsourcing. *International Journal of Hospitality Management,* In Press, Corrected Proof.

Lashley, C. (1998) Matching the management of human resources to service operations. . *International Journal of Contemporary Hospitality Management,* 10 (24), p. 24-33.

Lashley, C. (2008) Studying hospitality: insights from social sciences. *Scandinavian Journal of Hospitality & Tourism,* 8 (1), p. 69-84.

Lau, C.-M. and Ngo, H.-Y. (1996) One country many cultures: organizational cultures of firms of different country origins. *International Business Review,* 5 (5), p. 469-486.

Lau, C.-M. and Ngo, H.-Y. (2001) Organization development and firm performance: a comparison of multinational and local firms. *Journal of International Business Studies,* 32 (1), p. 95-114.

Lee, C. and Way, K. (2010) Individual employment characteristics of hotel employees that play a role in employee satisfaction and work retention. *International Journal of Hospitality Management,* 29 (3), p. 344-353.

Lee-Ross, D. and Pryce, J. (2010) *Human resources and tourism: skills, culture and industry.* Bristol, Channel View Publications.

Lewis, C. (2008) *How the East was won: the impact of multinational companies in Eastern Europe and the former Soviet Union (1989-2004).* London, Coptic.

Lim, D. H. and Wentling, R. M. (1998) Modification of transferred training programmes for multinational chain hotels in Korea. *International Journal of Training and Development,* 2 (1), p. 17-27.

Losekoot, E., et al. (2001) Conceptualising and operationalising the research interface between facilities management and hospitality management. *Facilities,* 19 (7/8), p. 296-303.

Lowell, B. (1999) *Race for the world : strategies to build a great global firm.* Boston, Harvard Business School Press.

Lu, Z. and Chiang, D. (2003) Strategic issues faced by Ontario hotels. *International Journal of Contemporary Hospitality Management,* 15 (6), p. 343-345.

Lucas, R. (2002) Fragments of HRM in hospitality? Evidence from the 1998 workplace employee relations survey. *International Journal of Contemporary Hospitality Management,* 14 (5), p. 207-212.

Lucas, R. and Deery, M. (2004) Significant developments and emerging issues in human resource management. *International Journal of Hospitality Management,* 23 (5), p. 459-472.

Magnini, V. P. and Honeycutt, E. D. (2003) Learning orientation and the hotel expatriate manager experience. *International Journal of Hospitality Management,* 22 (3), p. 267.

Maher, A. (1993) Measuring human resource value: an analysis based on the hotel industry. *Journal of Human Resource Costing & Accounting,* 1 (1), p. 15-33.

Marchante, A. J., et al. (2006) Determinants of skills shortages and hard-to-fill vacancies in the hospitality sector. *Tourism Management*, 27 (5), p. 791-802.

Martin, G. and Beaumont, P. (1998) Diffusing 'best practice' in multinational firms: prospects, practice and contestation. *The International Journal of Human Resource Management,* 9 (4), p. 671 - 695.

Mathews, V. E. (2000) Competition in the international hotel industry. *International Journal of Contemporary Hospitality Management,* 12 (2), p. 114-118.

Matias, Á., et al. (2009) An estimation of tourism dependence in French rural areas. *Advances in Tourism Economics,* p. 273-294.

Maxwell, G. and Lyle, G. (2002) Strategic HRM and business performance in the Hilton Group. *International Journal of Contemporary Hospitality Management,* 14 (5), p. 251-252.

Maxwell, G., et al. (2004) Quality service in the international hotel sector: A catalyst for strategic human resource development? *Journal of European Industrial Training,* 28 (2/3/4), p. 159-182.

Mayrhofer, W. and Brewster, C. (1996) In praise of ethnocentricity: expatriate policies in European multinationals. *International Executive,* 38 (6), p. 749-778.

McDowell, L., et al. (2007) Division, segmentation, and interpellation: the embodied labors of migrant workers in a Greater London hotel. *Economic Geography,* 83 (1), p. 1-25.

McEvoy, G. (1984) Small business personnel practices. *Journal of Small Business Management,* 23 (4), p. 1-8.

McGunnigle, P. J. and Jameson, S. M. (2000) HRM in UK hotels: a focus on commitment. *Employee Relations,* 22 (4), p. 403-422.

McPhail, R. and Fisher, R. (2008) It's more than wages: analysis of the impact of internal labour markets on the quality of jobs. *International Journal of Human Resource Management,* 19 (3), p. 461-472.

McSherry, B. (2006) The political economy of oil in Equatorial Guinea. *African Studies Quarterly,* 8 (3), p. 23-45.

Minbaeva, D., et al. (2003) MNC knowledge transfer, subsidiary absorptive capacity, and HRM. *Journal of International Business Studies,* 34 (6), p. 586-599.

Naama, A., et al. (2008) Human resource development issues for the hotel sector in Lybia: a government perspective. *International Journal of Tourism Research,* 10 (5), p. 481-492.

Nankervis, A. R. (2000) Human Resource Management strategies as competitive advantage: a case example of the hospitality sector in Southeast Asia & the Pacific Rim. *Research and Practice in Human Resource Management,* 8 (1), p. 111-133.

Ondrack, D. (1985) International transfers of managers in North American and European MNEs. *Journal of International Business Studies,* 16 (3), p. 1-19.

Orwell, G. (1933) *Down and out in Paris and London.* London, Penguin.

Paik, Y., et al. (2011) Interaction effects of globalization and institutional forces on international HRM practice: Illuminating the convergence-divergence debate. *Thunderbird International Business Review,* 53 (5), p. 647-659.

Pender, L. (2005) Managing the tourism system. *In*: Pender, L. and Sharpley, R. (eds.) *The management of tourism*. London, Sage, p. 1-13.

Perlmutter, H. V. (1969). The tortuous evolution of the multinational corporation. *Columbia Journal of World Business*, 4 (1), p. 9-18.

Pine, R. and Qi, P. (2004) Barriers to hotel chain development in China. *International Journal of Contemporary Hospitality Management,* 16 (1), p. 37-44.

Piso, A. (1999) Hotel and catering workers: class and unionisation. *Employee Relations,* 21 (2), p. 176-189.

Price, L. (1994) Poor personnel practice in the hotel and catering industry: does it matter? *Human Resource Management Journal,* 4 (4), p. 44-62.

Radolovic, O. (2010) Contractual liability of the guest for damage in direct hotelkeeper's contract. *Economic Research,* 23 (1), p. 97-114.

Redman, T. and Mathews, B. P. (1998) Service quality and human resource management: a review and research agenda. *Personnel Review,* 27 (1), p. 57-77.

Reisinger, Y. (2001) Unique characteristics of tourism, hospitality, and leisure services. *In*: Kandampully, J., et al. (eds.) *Service quality management in hospitality, tourism, and leisure*. Binghamton, Haworth Press, p. 15-50.

Riley, M. (1996) *Human resource management in the hospitality and tourism industry.* Oxford, Butterworth-Heinemann.

Riley, M., et al. (2002) *Tourism employment: analysis and planning.* Bristol, Channel View Publications.

Ritzer, G. (1998) *The McDonaldization thesis: explorations and extensions.* London, Sage.

Robson, P., et al. (1999) Low pay, labour market institutions, gender and part-time work: cross-national comparisons. *European Journal of Industrial Relations,* 5 (2), p. 187-207.

Rosenzweig, P. (2006) The dual logics behind international human resource management: pressures for global integration and local responsiveness. *In*: Björkman, I. and Stahl, G. (eds.) *Handbook of Research in International Human Resource Management*. Cheltenham, Edward Elgar, p.

Rosenzweig, P. M. and Nohria, N. (1994) Influences on human resource management practices in multinational corporations. *Journal of International Business Studies,* 25 (2), p. 229-251.

Rosenzweig, P. M. and Singh, J. V. (1991) Organizational environments and the multinational enterprise. *The Academy of Management Review,* 16 (2), p. 340-361.

Rowden, R. W. (1995) The role of human resource development in successful small to mid-sized manufacturing business: a comparative case study. *Human Resource Development Quarterly,* 6 (4), p. 355-373.

Rowley, G. and Purcell, K. (2001) 'As cooks go, she went': is labour churn inevitable? *International Journal of Hospitality Management,* 20 (2), p. 163-185.

Rutherford, D. G. and O'Fallon, M. J. (2007) *Hotel management and operations.* Hoboken, NJ, John Wiley & Sons.

Sartorius, K., et al. (2011) Human resource management and cultural diversity: a case study in Mozambique. *The International Journal of Human Resource Management,* 22 (9), p. 1963-1985.

Saunders, M., et al. (2007) *Research methods for business students.* Harlow, Pearson Education UK.

Schuler, R. S. and Rogovsky, N. (1998) Understanding compensation practice variations across firms: the impact of national culture. *Journal of International Business Studies,* 29 (1), p. 159-177.

Shamir, B. (1980) Between service and servility: role conflict in subordinate service roles. *Human Relations,* 33 (10), p. 741-756.

Shay, J. P. and Baack, S. A. (2004) Expatriate assignment, adjustment and effectiveness: an empirical examination of the big picture. *Journal of International Business Studies,* 35(3), p. 216-232.

Simons, T. and Enz, C. A. (1995) Motivating hotel employees. *Cornell Hotel and Restaurant Administration Quarterly,* 36 (1), p. 20-27.

Sisson, K., et al. (2003) All benchmarkers now? Benchmarking and the 'Europeanisation' of industrial relations. *Industrial Relations Journal,* 34 (1), p. 15-31.

Soltani, E. and Wilkinson, A. (2010) What is happening to flexible workers in the supply chain partnerships between hotel housekeeping departments and their partner employment agencies? *International Journal of Hospitality Management,* 29 (1), p. 108-119.

Stake, R. (2000). Case studies. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 435-454). Thousand Oaks: Sage.

Szulanski, G. (1996) Exploring internal stickiness: impediments to the transfer of best practice within the firm. *Strategic Management Journal,* 17, p. 27-43.

Tanke, M. L. (2001) *Human resource management for the hospitality industry.* Albany, Delmar.

Taylor, S., et al. (1996) Toward an integrative model of strategic international human resource management. *The Academy of Management Review,* 21 (4), p. 959-985.

Teare, R. (1995) The international hospitality business: a thematic perspective. *International Journal of Contemporary Hospitality Management,* 7 (7), p. 55 - 73.

Timo, N. and Davidson, M. (2005) A survey of employee relations practices and demographics of MNC chain and domestic luxury hotels in Australia. *Employee Relations,* 27 (2), p. 175-192.

UNCTAD (2004) *World investment report 2004: the shift towards services*. [online]. Available from:http://www.unctad.org/Templates/WebFlyer.asp?intItemID= 3235&lang =1 [Accessed 17th September 2009].

Walsh, T. J. (1990) Flexible labour utilisation in the private service sector. *Work, Employment & Society,* 4 (4), p. 517-530.

Welch, C., Piekkari, R., Plakoyiannaki, E., & Paavilainen-Mantymaki, E. (2011). Theorising from case studies: Towards a pluralist future for international business research. *Journal of International Business Studies, 42*(5), 740–762.

Whitla, P., et al. (2007) Global strategies in the international hotel industry. *International Journal of Hospitality Management,* 26 (4), p. 777-792.

Whitley, R. (1999). *Divergent capitalisms: The social structuring and change of business systems*. Oxford: Oxford University Press, (Chapter 2).

Wilton, N. (2006) Strategic choice and organisational context in HRM in the UK hotel sector. *Service Industries Journal,* 26 (8), p. 903-919.

Wood, G. (2004) Business and politics in a criminal state: the case of Equatorial Guinea. *African Affairs,* 103 (413), p. 547-567.

Wood, R. C. (1992) *Working in hotels and catering*. London, Routledge.

Woods, R. H. (1999) Predicting is difficult, especially about the future: human resources in the new millennium. *International Journal of Hospitality Management,* 18 (4), p. 443-456.

Woodworth, R. M. and Mandelbaum, R. (2010) Seventy-five years of U.S. hotel revenues, expenses, and profits. *Cornell Hospitality Quarterly,* 51 (1), p. 20-26.

Wright, P. M., et al. (2005) New models of strategic HRM in a global context. *The International Journal of Human Resource Management,* 16 (6), p. 875-881.

WTO (2001) *Tourism 2020 Vision: global forecasts and profiles of market segments.* Madrid, WTO.

WTO (2011) *International tourism 2010: multi-speed recovery.* [online]. Available from: http://85.62.13.114/media/news/en/press_det.php?id=7331&idioma=E [Accessed 27/06/2011].

WTTC (2011) *World -key facts at a glance*. [online]. Available from: http://www.wttc.org/eng/Tourism_Research/Economic_Research/ [Accessed 17/02/2011].

Yan, Y., et al. (2007) Vertical integration of corporate management in international firms: implementation of HRM and the asset specificities of firms in China. *International Journal of Human Resource Management,* 18 (5), p. 788-807.

Yeung, A. (2006) Setting people up for success: how the Portman Ritz-Carlton hotel gets the best from its people. *Human Resource Management,* 45 (2), p. 267-275.

Yu, Y., Byun, W.-H., & Lee, T. J. (2014). Critical issues of globalisation in the international hotel industry. *Current Issues in Tourism, 17*(2), 114-118.

Zhang, H. Q. and Wu, E. (2004) Human resources issues facing the hotel and travel industry in China. *International Journal of Contemporary Hospitality Management,* 16 (7), p. 424-428.

Zuehl, J. J. and Sherwyn, D. S. (2001) Identifying uniform employment termination practices for multinational employers. *Cornell Hotel and Restaurant Administration Quarterly,* 42 (5), p. 72-85.

Appendix

The seven properties fully owned by HotelCo participating to the survey are located in:

- CONTINENTAL EUROPE
  - United Kingdom (1 hotel)
  - Switzerland (1 hotel)
  - Germany (3 hotels)

- CAUCASUS
  - Azerbaijan (1 hotel)

- CENTRAL ASIA
  - Kyrgyzstan (1 hotel)

Examining the dynamics and strategies of entrepreneurial families: Lessons from Mexico

Mariana Estrada-Robles

Doctoral Researcher. Management School, University of Sheffield, UK,
`m.estradarobles@sheffield.ac.uk`

Abstract. Entrepreneurship and family business have until recently been developed as two independent fields, leaving academics with the opportunity to explore their overlap and intersection. However, there is still scarce knowledge about how some families can act entrepreneurially. The present paper, as a research proposal, intends to reduce that gap by focusing on entrepreneurial families; explicitly, families with more than one entrepreneur and/or business. The purpose of the study is to provide an explanation on the way family dynamics and family members' actions materialize into entrepreneurial activity. Multiple case studies are to be conducted in Mexico, Latin American country where family firms are predominant and represent main contributors to job and wealth creation. This research is among few studies that address entrepreneurial families under that cultural and economic context. Findings intend to show family links to entrepreneurship, theory might be developed to further explore those connections providing academics in entrepreneurship and family business with ample opportunities for future research.

Keywords: *entrepreneurship; family business; entrepreneurial orientation; entrepreneurial family; portfolio entrepreneurship.*

1   Introduction

Family business and entrepreneurship have until recently been studied separately. Generally, entrepreneurship has been concerned with the processes of discovery, evaluation and exploitation of new business opportunities (Shane and Venkataraman, 2000), while family business research has been more focused on continuity and sustainability. Until now, the role that family plays on entrepreneurship is still not fully understood (Cruz and Nordqvist, 2012).

Family businesses constitute large contributors to world economy impacting positively in wealth and job creation (Astrachan and Shanker, 2003, Gersick et al., 1997). In Latin American countries for example, family firms count as the majority of businesses, while in several European economies they constitute at least 60% of private-owned firms (IFERA, 2003). Although family firms represent the vast majority of

business around the world, they are always facing internal and external challenges. Renewal capabilities to regenerate the business through corporate entrepreneurship or through new venture creation are needed for long-term survival (Zahra et al., 2004, Cruz and Nordqvist, 2012). These entrepreneurial capabilities may be characterized by: the specific resources distinctive of a family in a firm (Habbershon et al., 2003); cultural patterns (Zahra et al., 2004); and family lifecycle, firm stage and level of family involvement (Hall et al., 2001).

Family member participation in business is very common in emerging economies and collectivist cultures. Members may not only come from the nuclear but the extended family as well. The existence of numerous members in families can lead to the creation of business groups for distinct reasons beyond financial objectives (Carter and Ram, 2003). Family relatives help other relatives to create new firms, forming a portfolio of business in the family, in other words, members acting together in the pursuit of opportunity in the form of family entrepreneurial teams. (Discua Cruz et al., 2013). Nonetheless, general understanding in entrepreneurial teams is increasing (Forbes et al., 2006, Francis and Sandberg, 2000), very little has been done in entrepreneurial teams formed by family members (Discua Cruz et al., 2013).

The stream of literature in family business in Latin America has typically focused on succession, governance or strategy. Until recently, studies have focused on the contribution of family businesses to entrepreneurship, for example the Global Succesful Transgenerational Entrepreneurship Practices (STEP) Project focuses in entrepreneurial orientation across generations (Nordqvist et al., 2011, Cruz and Nordqvist, 2012; Global STEP Project); however, it centres more in multigenerational processes rather than in understanding the family dynamics and strategies of the multiple entrepreneurs within a family with different businesses.

In this study 'entrepreneurial families' are termed as families with more than one entrepreneur bonded by blood or marriage that own/manage more than one business, and show entrepreneurial orientation in the recognition and exploitation of business opportunities. These type of families may or not be working together in formal family entrepreneurial teams.

Case studies will be analysed and they will be based in Mexico. As one of the largest countries in Latin America, Mexico provides a rich context for family business research in the following ways: family businesses constitute the great majority of private firms in Mexican economy; culture is considered a collectivist culture in which family and business issues are practically inseparable; and business groups are common and owned by a single family. Moreover, one of the challenges faced by Latin American and Mexican family firms is the uncertainty in the institutional, economic and political environment alongside internal familial issues. This can provide the research with the possibility of making it more valuable.

The present study will contribute in four areas. First, it builds on recent research that links entrepreneurship and family business together, understanding the influence that family exerts on entrepreneurial activities. Second, it adds to the scarce literature in entrepreneurial orientation present in family members that engage in new venture creation. Third, it provides an overview into how entrepreneurial orientation is materialized into everyday life activities. Fourth, it provides a different perspective in port-

folio entrepreneurship by having entrepreneurial families as focus of study. Finally, all these contributions will also add to the extent literature on entrepreneurship and family business in Latin America and Mexico.

This paper is organized as follows. The theoretical platform is initially presented by addressing definitions of family firms, and exploring entrepreneurial orientation and entrepreneurial families. At the end of that section, the research question and purpose of the study is stated. Later, the focus of research is briefly introduced with a review of Mexico and family businesses. The research methodology is then proposed for the specific research purpose. Finally, a conclusion is drawn with a discussion on the relevance of the research and potential implications for future work.

2 Literature Review

An overview of family firms

A family as a unit of analysis for research is not without complication(Nordqvist and Melin, 2010). Especially because of the dynamism of this social institution through time and its changing definition (Koerner and Fitzpatrick, 2002). Currently, the perspective of families is different, views and behaviours are likely to be influenced by a combination of the surrounding culture and family values, as families migrate across nations (Sharma and Manikutty, 2005). Despite the fact that family, culture and business are highly connected, researchers agree that study is still insufficient and that the field has clear potential (Westhead and Cowling, 1998, Sharma, 2004).

Family enterprise is regarded as a combination and overlap of business and family (Basco and Rodriguez, 2009). Because its heterogeneity and complexity, there is a lack of consensus surrounding the definition of family business (Sharma, 2004, Anderson et al., 2005). Most efforts to define family firms focus on differentiating them from nonfamily firms (Sharma, 2004, Gersick et al., 1997, Tagiuri and Davis, 1996). Researchers have used different criteria to define these businesses such as ownership, control, multi-generation involvement and intention for the business to continue in the family (Astrachan, 2003, Westhead and Cowling, 1998).

One of the most common used approaches is the one developed by Westhead and Cowling (1998) that focuses on ownership by a family group related by blood or marriage with self-perception of the firm being a family business. Chua et al. (1999) add successive generations and sustainability, they claim that a family business is the one managed and/or governed by one or a few families with the intention to shape and pursue the vision in a way that is potentially sustainable across generations in the family core. Some family businesses go further on this and evolve to become 'entrepreneurial families'.

From Entrepreneurial Orientation to Entrepreneurial Families

Nordqvist and Melin (2010) recognize that family often helps any entrepreneur for new venture creation within and beyond the family's boundaries, acknowledging family influences in any business. However, researchers had paid little attention to this social institution as an influencer on the emergence of new business ventures (Aldrich and Cliff, 2003).

Not long ago entrepreneurship begun to acknowledge the family embeddedness perspective (Anderson et al., 2005, Aldrich and Cliff, 2003); the exploration of the linkages between entrepreneurial activity and the family dimension (Heck et al., 2008); and the entrepreneurial orientation (Cruz and Nordqvist, 2012, Rauch et al., 2009, Short et al., 2009, Sieger et al., 2011, Naldi et al., 2007).

Entrepreneurial orientation (EO) regards to the "the organizational processes, methods and styles that firms use to act entrepreneurially" (Lumpkin and Dess, 1996); accordingly, the capabilities that an entrepreneur shows to respond to external opportunities (Martinez and Aldrich, 2014).

Entrepreneurial orientation is considered a framework to address corporate entrepreneurship, from both perspectives: innovating within an organization or through the creation of new organizations. Lumpkin and Dess (1996) referred EO as the processes, practices and decision-making activities that lead to new entry. On the other hand, Miller (1983) proposed the following dimensions to assess EO in an organization: innovativeness, risk taking, proactiveness, autonomy and competitive aggressiveness.

In family context, EO and the role of family is still not fully understood (Cruz and Nordqvist, 2012). Most research lies in identifying if family firms provide a scenario that stimulates or prevents the firm to act entrepreneurially; nonetheless, researchers have not yet arrived to an agreement (Zahra, 2005, Short et al., 2009, Naldi et al., 2007). On one side, studies support the idea that the family firm setting fosters entrepreneurial activities while taking advantage of their distinctive characteristics such as long-term orientation, pursuit of financial and non-financial goals, and established family relationships (Aldrich and Cliff, 2003, Rogoff and Heck, 2003). On the other side, studies highlight that family firms become conservative and resistant to change beyond the founding entrepreneurial generation (Cabrera-Suárez et al., 2001, Ward, 1987), hence inhibiting EO (Zahra, 2005).

Taking this into consideration, an entrepreneurial family is defined by Nordqvist and Melin (2010) as a social institution that can drive or constrain entrepreneurial activity; influenced by the family's access to resources (e.g. financial and social capital) and the family's specific culture (e.g. norms, attitudes and values). However, the basic premise of entrepreneurship about discovery and exploitation of new business opportunities is not latent in the latter definition; nor that EO can provide family members with the possibility to engage in entrepreneurial processes related to founding or acquiring multiple businesses. Consequently, for this study entrepreneurial family is defined as a social entity with more than one entrepreneur bonded by blood or marriage that own/manage more than one business; and show entrepreneurial orientation in the recognition and exploitation of business

opportunities. As opposed to a family business which focuses more on continuity of the existing firm.

Entrepreneurial families form business groups as they engage in portfolio entrepreneurship. The reasons behind portfolio entrepreneurship may be diverse considering: lateral growth strategy (Rosa and Scott, 1999); wealth and risk diversification (Sieger et al., 2011); long-term orientation (Miller and Le Breton-Miller, 2005); opportunity development for the next generations or succession solving (Carter and Ram, 2003). Entrepreneurial teams may exist in entrepreneurial families to create/acquire a new business or acting together in existing organizations (Discua Cruz et al., 2013). Normally, portfolio entrepreneurship has been studied from an individual level (habitual entrepreneur) rather than from a collective level (Ucbasaran et al., 2003, Westhead and Wright, 1998). In family business context, business groups or family portfolio entrepreneurship should have the family as unit of analysis, in specific the entrepeneurial family whose members might act as a team.

Studies propose that research should go beyond the focus of family business but instead concentrate on the family team behind entrepreneurial processes (Chrisman et al., 2003, Ucbasaran et al., 2003). Discua Cruz et al. (2012) suggest that further evidence is needed revolving around the operation patterns and development of these patterns in family entrepreneurial teams. Other academics mention the need to study portfolio entrepreneurship not only from the founder perspective but also from the members of the family that constitute the team, including as well the interconnection of the established ventures (Iacobucci and Rosa, 2005).

Because the level of entrepreneurship in a business may be influenced by external and internal context (Lumpkin and Dess, 1996), culture should also be addressed in this perspective especially because it is more significant to EO for family firms than non-family firms (Zahra et al., 2004). Research has shown that cultural patterns of family business can either facilitate or constrain the level of corporate entrepreneurship, depending on the number of members involved, the openness of the culture and degree of explicitness (Hall et al., 2001). For example, family firms with their unique organizational context can develop entrepreneurship displayed in the form of new products, ventures and ideas (Naldi et al., 2007).

To make a link between family, business and entrepreneurial activity where multiple entrepreneurs coexist, this research sets 'entrepreneurial families' as focus of study. Drawing upon that, the central question in this study asks **how do family dynamics work in an entrepreneurial family, analysing the way members influence each other in their everyday life for the execution of entrepreneurial activity?**

The answer, however, does not rely only on business strategies but on day-to-day activities and practices that in turn are affected by family dynamics. The focus is placed on highlighting the variety of relational dynamics, analysing everyday life actions of the multiple entrepreneurs within the family that in turn enable or hinder entrepreneurial activity throughout the different businesses in the family.

Shifting the emphasis from the business, the centre of this research is placed in the family with an entrepreneurial perspective; moving away from the traditional family business topics such as succession, ownership, growth or performance to an in-depth

examination of entrepreneurship in the family context. Furthermore, it examines multiple entrepreneurs within a family rather than considering a sole entrepreneur. It also analyses the influence of the members in the family unit that plays a central role in entrepreneurship processes, going a step further in examining family dynamics. All in all, this research aims to provide a nuanced research that is able to grasp the complexity of the dynamics between the individuals that make up an entrepreneurial family through the analysis of distinct formations of families and businesses in a specific cultural and institutional context.

3   Context of the study

Mexico is one of the largest countries in Latin America with a population of c. 115 million inhabitants (INEGI, 2013). Most family firms in the country are SMEs, in general they have great importance in Mexican economy because they represent 99.8% of all businesses (GEM, 2010). About 90% of family business in Mexico generate 70% of national employment and constitute 60% of GDP (Secretaría de Economía del Gobierno Federal Mexicano, , 2011). They concentrate mainly in the industrial, financial and wholesale sectors; others falling in construction, retail selling and hospitality sectors[1].

Family businesses normally start as entrepreneurial companies and then become family-owned businesses, showing high ownership concentration and substantial family control. As Steier (2009) states, family ownership is the major form of governance in emerging economies. In Mexico, family firms are owned and managed by either one or a group of families, or by successive generations of the founding family (Aguiló and Aguiló, 2012).

Some characteristics in Latin American countries are common, most of them sharing histories, languages and culture (Discua Cruz et al., 2013). According to Gupta and Levenburg (2010), it is common to have family and business intertwined in Latin American countries, like in Mexico. In fact, the same authors argue that the separation of both entities becomes more complex in collectivist cultures. People in these cultures value social networks and relationships; and success and wealth are shared among family and friends (Hofstede, 1984). In order to do business, a developed relationship with the other person is required to cultivate trust; reason why normally businesses are ran by family members (Gupta and Levenburg, 2010). The preference for family members goes beyond in emerging economies because there is a greater dependence on personal trust than on institutional trust (Khanna and Rivkin, 2001). Moreover, due to the level of uncertainty for business in these developing countries, family firms strongly rely on family ties, social networks and relationships (Nordqvist et al., 2011). In addition to financial objectives, family firms in Latin American coun-

---

[1] During the second semester of 2012, PricewaterhouseCoopers conducted the Global Family Business Survey in over 30 countries including Mexico. Encuesta de Empresas Familiares 2012. *pwc.com*. Retrieved March 20, 2014 from http://www.pwc.com/mx/es/encuesta-empresas-familiares/archivo/2013-01-empresa-familiar-finales.pdf

tries also aim to achieve nonfinancial goals such as: family prestige; sustainability; and reputation among family members, employees and other contacts (Gupta and Levenburg, 2010).

It is common in Latin America and in Mexico to find portfolio entrepreneurship. Business groups are more prominent in developing countries because they cultivate unique advantages in the particular social, economic and political conditions of these countries (Khanna, 2000); correspondingly, entrepreneurial families are more numerous in emerging economies. Entrepreneurial families coming from a collectivist culture rely on relatives collaboration and some families act as entrepreneurial teams (Discua Cruz et al., 2013). The search of new business opportunities is evident when families are numerous; in Mexico, family members with strong ties normally go beyond the nuclear family.

It is important to highlight that Mexico recently went through one the greatest economic crunches in its history: the crisis that started in the United States by the end of 2007, and then spread as a global crisis. Mexico was one of the most severely affected countries as a result of its high dependence to the US in terms of trade and investment. All types of firms were affected, leaving SMEs and family business with a great challenge to overcome the reduction in overall sales and low levels of liquidity. This economic event should not remain unnoticed in the research of Mexican family businesses.

Although family businesses perceive they are valued by the Mexican government (PwC, 2013), they also recognize that more support from the government is needed. This is more important especially now that Mexico is living a period of transition because of the change of President in late 2012. With that event, major reforms regarding financial, energetic, labour and education sectors were passed in 2013; commencing to be effective from 2014 onwards. It would be of considerable significance to discover if these changes are significant in family business context. Relevance of the study of family firms in Mexico should never be overlooked as they have played an essential role in the history, and the economic, social and political development of the country.

4 Proposed Methodology

The study will rely on in-depth qualitative case studies based in Mexico. Multiple case study design is helpful in the precision and validity of the result to develop theoretical insight (Neergaard, 2007), which arises from methodological rigor and multi-case comparative logic (Eisenhardt, 1991). Case studies are the most convenient to answer 'how' and 'why' questions (Yin, 1994) under an exploratory study that will help understand the complexity and nuances of the subject to answer the research question.

In regards to the guidelines for appropriate sampling in entrepreneurship research (Neergaard, 2007) and the particular objective of the study, purposeful sampling will be used for case selection. Knowing that family business are heterogeneous entities

(Handler, 1989a, Howorth et al., 2010, Westhead and Howorth, 2007), the most suitable sampling strategy is 'maximum variation' combined with criterion sampling (Neergaard, 2007). With this combination not only data collection and analysis will provide in-depth description of each case, but also common characteristics and recognized patterns will increase the robustness of the findings (Neergaard, 2007).

The population for the sample will come from the support of the Entrepreneurial Families Centres Network from the Tecnologico y de Monterrey – the largest university in Latin America. The extent of the Network, comprised of 17 regional Centres for Family Enterprise will give the study the opportunity and flexibility to find the right family profiles. The requirement for the entrepreneurial families is to have at least two members managing/owning at least one entrepreneurial venture.

The suitability of the sample size is relative in qualitative research (Sandelowski et al., 1992). According to Eisenhardt (1989), it is challenging to create theory form less than 4 cases; however, more than 10 becomes difficult to manage and analyse. The decision of the number of cases selected will be made by the sufficiency and quality of the data obtained and whether the sample facilitates the understanding of the research process (Neergaard, 2007). It will also highly depend on the availability of the families and how much each one adds incrementally (Eisenhardt, 1991).

Semi-structured interviews will be conducted. An initial interview guideline will be designed. The guideline will capture in general the following aspects: everyday life activities, recognition of opportunities, family members' influence over other members, factors surrounding family businesses, and internal and external challenges. Under a context of discovery, this study will adopt an inductive approach. Conducting data collection and analysis will come simultaneously in an iterative process (Denzin and Lincoln, 2011). Thus, repeated interviews to family members will be required moving from data to emerging concepts, and backwards.

Interviews will be recorded and transcribed. As respondents will answer in the official language in Mexico, Spanish, the transcription process will be followed by English translation carried out by the researcher.

Depending on the size of the entrepreneurial families, different members will be interviewed. Normally, men predominate as business founders in Mexico; however, in these case studies diversity of members will be considered. Old and young generation, men and women will be interviewed, depending on the profiles of family members in family businesses. With these, the known limitation of single informants to draw conclusions on entrepreneurial subjects is avoided (Sharma, 2004, Short et al., 2009, Nordqvist and Melin, 2010). Interviews to non-family members and stakeholders will be useful to achieve reliability of findings.

Entrepreneurial family members may be in pursuit of new opportunities at the time of the study or may have already formed their portfolio of business; in the latter circumstance, data will adopt a retrospective angle. Thus, flawed memory is acknowledged in this process as a potential limitation. However, this is another reason why multiple respondents (family members, non-family members and stakeholders) will be considered for the interviews to capture the evolution of the story and reduce the limitation.

In regards to the analysis stage, previous research with multi-case study methodology is taking into consideration (Nordqvist and Melin, 2010, Zellweger and Sieger, 2012). No coding software will be used in order to make a more profound analysis of interview data (Discua Cruz et al., 2013). Reports per each case study will be developed under a common template following the interview guideline and organised by topics. Reports allow researchers to easily identify patterns and constructs without the use of a coding software (Eisenhardt, 1989). For triangulation processes other sources of evidence such as secondary data can be useful for the researcher to get a wider picture of the family, the businesses and the members' activities both in the family and in the firm. With comparative cross-case analysis the development of topics and theorizations is possible (Discua Cruz et al., 2013); hence, frequencies, differences and patterns will be determined thanks to cross-case analysis.

The general research process involving field notes, secondary data and case study reports will provide the opportunity to assure the quality of the study from data collection to the reliability of outcomes. To achieve validity, the iterative process will help alongside with giving respondents the opportunity to know the empirical material and make comments on it, aiming for a high level of accuracy and consensus (Lincoln and Guba, 1985).

As this research will study entrepreneurial families over time involving the collection of empirical data, methodological challenges may arise such as ethical implications. The investigation of everyday actions in the family and in the firm to discover entrepreneurial activity may bring up sensitive topics for both the family members and the firms, resulting in ethical considerations for the researcher. Thus, confidentiality and anonymity should be considered.

After the evaluation of the proposed methodology, case studies will provide the understandings to achieve the aim of the study of examining how family dynamics and family members' strategies materialize into entrepreneurial activity in entrepreneurial families.

5 Conclusions

With reference to the main stream of literature in entrepreneurship and family businesses, and its continuous request to link both fields; the current study intends to reduce that gap by focusing on entrepreneurial families.

The conclusions drawn will not only add to the international literature on entrepreneurship and family businesses but also to the extent studies in Latin American contexts. The study will provide a useful background for future research in Mexico and in other countries sharing common characteristics.

The exploration of family interactions and its members through case study research, will build to the understanding of entrepreneurial activity reflected in everyday actions. Crumbling the complexity of family dynamics will make the contribution to knowledge on the nature of entrepreneurial families.

Potential implications may be drawn by the resulting findings. This in-depth study can provide a call to motivate entrepreneurial orientation in day-to-day family business practice for the emergence of entrepreneurial families; thus, capitalizing their capabilities as job and wealth creators. It can also elaborate on a clearer vision of the nature of family business for policy makers to develop special support for them. Finally, this research intends to encourage the emergence of new research paths from entrepreneurship and family business academics.

References

Aguiló, Tomas & Aguiló, Nicolas (2012) Family business performance: evidence from Mexico. *Cuadernos de Administración,* 25**,** 39-61.

Aldrich, Howard E. and Cliff, Jennifer. E. (2003) The pervasive effects of family on entrepreneurship: toward a family embeddedness perspective. *Journal of Business Venturing,* 18**,** 573-596.

Anderson, Alistair R. et al. (2005) The role of family members in entrepreneurial networks: Beyond the boundaries of the family firm. *Family Business Review,* 18**,** 135-154.

Astrachan, Joseph H. (2003) Commentary on the special issue: the emergence of a field. *Journal of Business Venturing,* 18**,** 567-572.

Astrachan, Joseph H. and Shanker, Melissa C. (2003) Family Businesses' Contribution to the U.S. Economy: A Closer Look. *Family Business Review,* 16**,** 211-219.

Basco, Rodrigo and Perez-Rodriguez, Maria J. (2009) Studying the Family Enterprise Holistically Evidence for Integrated Family and Business Systems. *Family Business Review,* 22**,** 82-95.

Cabrera-Suárez, Katiuska. et al. (2001) The Succession Process from a Resource- and Knowledge-Based View of the Family Firm. *Family Business Review,* 14**,** 37-46.

Carter, Sara and Ram, Monder (2003) Reassessing portfolio entrepreneurship. *Small Business Economics,* 21**,** 371-380.

Chrisman, James J. et al. (2003) An introduction to theories of family business. *Journal of Business Venturing,* 18**,** 441-448.

Chua, Jess H. et al. (1999) Defining the family business by behavior. *Entrepreneurship, Theory and Practice,* 23**,** 9-39.

Cruz, Cristina and Nordqvist, Mattias (2012) Entrepreneurial orientation in family firms: a generational perspective. *Small Business Economics,* 38**,** 33-49.

Denzin, Norman and Lincoln, Yvonna S. (2011) *The Sage Handbook of Qualitative Research,* Thousand Oaks, CA, Sage.

Discua Cruz, Allan et al. (2012) Understanding entrepreneurial cultures in family businesses: A study of family entrepreneurial teams in Honduras. *Journal of Family Business Strategy,* 3**,** 147-161.

Discua Cruz, Allan et al. (2013) Intrafamily Entrepreneurship: The Formation and Membership of Family Entrepreneurial Teams. *Entrepreneurship Theory and Practice,* 37**,** 17-46.

Economia, Secretaria. (2011) Las PyMES generan siete de cada 10 empleos formales en el país. Mexico.

Eisenhardt, K. M. (1989) Building theories from case study research. *Academy of management review,* 14**,** 532-550.

Eisenhardt, K. M. (1991) Better stories and better constructs: the case for rigor and comparative logic. *Academy of Management review,* 16**,** 620-627.

Forbes, Daniel P. et al. (2006) Entrepreneurial team formation: An exploration of new member addition. *Entrepreneurship Theory and Practice,* 30**,** 225-248.

Francis, Deborah H. and Sandberg, William R. (2000) Friendship Within Entrepreneurial Teams and its Association with Team and Venture Performance. *Entrepreneurship Theory and Practice,* 25**,** 5–26.

GEM (2010) National Summaries. Global Entrepreneurship Monitor.

Gersick, Kelin E. et al. (1997) *Generation to Generation: Life Cycles of the Family Business,* Cambridge, MA, Harvard Business School Press.

Gupta, Vipin and Levenburg, Nancy (2010) A Thematic Analysis of Cultural Variations in Family Businesses: The CASE Project. *Family Business Review,* 23**,** 155-169.

Habbershon, Timothy G. et al. (2003) A unified systems perspective of family firm performance. *Journal of Business Venturing,* 18**,** 451-465.

Hall, Anika et al. (2001) Entrepreneurship as Radical Change in the Family Business: Exploring the Role of Cultural Patterns. *Family Business Review,* 14**,** 193-208.

Handler, W. C. (1989) Methodological Issues and Considerations in Studying Family Businesses. *Family Business Review,* 3**,** 257-276.

Heck, Ramona K. Z. et al. (2008) Emerging paths of family entrepreneurship research. *Journal of Small Business Management,* 46**,** 317-330.

Hofstede, G. (1984) The Cultural Relativity of the Quality of Life Concept. *Academy of Management Review,* 9**,** 389-398.

Howorth, Carole et al. (2010) Family firm diversity and development: An introduction. *International Small Business Journal,* 28**,** 437-451.

Iacobucci, Donato and Rosa, Peter (2005) Growth, diversification, and business group formation in entrepreneurial firms. *Small Business Economics,* 25**,** 65-82.

IFERA (2003) Family businesses dominate: International family enterprise research academy (IFERA). *Family Business Review,* 4**,** Family Business Review.

INEGI (2013) Proyeccion de Poblacion Mexico 2013. Anuario estadístico y geográfico de los Estados Unidos Mexicanos 2013. Aguascalientes, Mexico, Instituto Nacional de Estadística y Geográfia.

Khanna, T. (2000) Business groups and social welfare in emerging markets: Existing evidence and unanswered questions. *European Economic Review,* 44**,** 748-761.

Khanna, Tarun and Rivkin, Jan W. (2001) Estimating the performance effects of business groups in emerging markets. *Strategic management journal,* 22**,** 45-74.

Koerner, Ascan F. and Fitzpatrick, Mary Anne (2002) Toward a theory of family communication. *Communication Theory,* 12**,** 70-91.

Lincoln, Yvonna S. and Guba, Egon G. (1985) *Naturalistic inquiry,* Beverly Hills, CA, Sage.

Lumpkin, G. T. and Dess, Gregory G. (1996) Clarifying the entrepreneurial orientation construct and linking it to performance. *Academy of Management Review,* 21**,** 135-172.

Martinez, Martha A. and Aldrich, Howard E. (2014) Sociological Theories Applied to Family Businesses. *In:* L. Melin *et al.* (Eds.) *The SAGE Handbook of Family Business.* London, SAGE Publications Ltd.

Miller, Danny (1983) The Correlates of Entrepreneurship in three types of firms. *Management Science,* 29**,** 770-791.

Miller, Danny and Le Breton-Miller, Isabelle (2005) *Managing for the long run: Lessons in competitive advantage from great family businesses.,* Boston, Harvard Business School Press.

Naldi, Lucia et al. (2007) Entrepreneurial orientation, risk taking, and performance in family firms. *Family Business Review,* 20**,** 33-47.

Neergaard, H. (2007) Sampling in entrepreneurial settings. *In:* H. Neergaard *et al.* (Eds.) *Handbook of qualitative research methods in entrepreneurship.* Cheltenham, U.K, Edward Elgar Publishing.

Nordqvist, Mattias et al. (2011) Understanding entrepreneurial family businesses in uncertain environments: the case of Latin America. In M. Nordqvist *et al.* (Eds.) *Understanding Entrepreneurial Family Businesses in Uncertain Environments: Opportunities and Resources in Latin America.* Cheltenham, Edward Elgar Publishing Ltd.

Nordqvist, Mattias and Melin, Leif. (2010) Entrepreneurial families and family firms. *Entrepreneurship and Regional Development,* 22**,** 211-239.

Rauch, Andreas et al. (2009) Entrepreneurial Orientation and Business Performance: An Assessment of Past Research and Suggestions for the Future. *Entrepreneurship Theory and Practice,* 33**,** 761-787.

Rogoff, Edward G. and Heck, Ramona K. Z. (2003) Evolving research in entrepreneurship and family business: recognizing family as the oxygen that feeds the fire of entrepreneurship. *Journal of Business Venturing,* 18**,** 559-566.

Rosa, Peter and Scott, Michael (1999) The prevalence of multiple owners and directors in the SME sector: implications for our understanding of start-up and growth. *Entrepreneurship & Regional Development,* 11**,** 21-37.

Sandelowski, Margarete et al. (1992) Using qualitative and quantitative methods: The transition to parenthood of infertile couples. *Qualitative methods in family research***,** 301-322.

Shane, Scott and Venkataraman, S. (2000) The promise of entrepreneurship as a field of research. *Academy of Management Review,* 25**,** 217-226.

Sharma, P. (2004) An Overview of the Field of Family Business Studies: Current Status and Directions for the Future. *Family Business Review,* 17**,** 1-36.

Sharma, Pramodita and Manikutty, S. (2005) Strategic divestments in family firms: Role of family structure and community culture. *Entrepreneurship Theory and Practice,* 29**,** 293-311.

Short, Jeremy C. et al. (2009) Family Firms and Entrepreneurial Orientation in Publicly Traded Firms A Comparative Analysis of the S&P 500. *Family Business Review,* 22**,** 9-24.

Sieger, Phillip et al. (2011) Portfolio entrepreneurship in family firms: a resource-based perspective. *Strategic Entrepreneurship Journal,* 5**,** 327-351.

Steier, Lloyd P. (2009) Familial capitalism in global institutional contexts: Implications for corporate governance and entrepreneurship in East Asia. *Asia Pacific Journal of Management,* 26**,** 513-535.

Tagiuri, Renato and Davis, John (1996) Bivalent Attributes of the Family Firm. *Family Business Review,* 9**,** 199-208.

Ucbasaran, Deniz et al. (2003) A longitudinal study of habitual entrepreneurs: starters and acquirers. *Entrepreneurship and Regional Development,* 15**,** 207-228.

Ward, J. L. (1987) *Keeping the family business healthy,* San Francisco, Jossey-Bass. Westhead, Paul and Cowling, Marc. (1998) Family firm research: The need for a methodology to rethink. *Entrepreneurship Theory and Practice,* 23**,** 31-56.

Westhead, Paul and Howorth, Carole (2007) Types' of private family firms: An exploratory conceptual and empirical analysis. *Entrepreneurship and Regional Development,* 19**,** 405-431.

Westhead, Paul and Wright, Mike (1998) Novice, portfolio, and serial founders: Are they different? *Journal of Business Venturing,* 13**,** 173-204.

Yin, R. (1994) *Case Study Research. Design and Methods.,* Thousands Oaks, CA., Sage.

Zahra, Shaker A. (2005) Entrepreneurial risk taking in family firms. *Family Business Review,* 18**,** 23-40.

Zahra, Shaker A. et al. (2004) Entrepreneurship in family vs. non-family firms: A resource-based analysis of the effect of organizational culture. *Entrepreneurship Theory and Practice,* 28**,** 363-381.

Zellweger, Thomas M. and Sieger, Phillip (2012) Entrepreneurial orientation in long-lived family firms. *Small Business Economics,* 38**,** 67-84.

Challenges Faced by Entrepreneurs and Owners of Greek SMEs in a Crisis Context

Joanna Konstantinou

Management School Sheffield University,
`jkonstantinou@seerc.org`

Abstract. The aim of this study is to explore the financial challenges faced by entrepreneurs and owners of Greek small sized SMEs, whose businesses are currently operating in the context of financial crisis. Within a hermeneutic phenomenological perspective, semi-structured interviews were conducted with two individuals and the focus was to identify which, if any, financial challenges currently faced by their businesses can be attributed to the crisis. The findings were grouped with the goal of identifying typology of challenges faced, the types of challenges attributed to financial crisis and the functional role of attributions. Further research is suggested in order to identify which of these experienced challenges can be attributed to the leadership style of the entrepreneurs and which are the strategies that they are implementing in order to acquire resilience.

Keywords: *SMEs, Challenges, Crisis, Entrepreneurship*

Introduction

This study has explored the findings of other researchers in identifying the financial challenges and problems of Small to Medium sized Enterprises (SMEs) around the world. Indeed, literature review was carried out in two phases: in the first phase the literature review had the focus of identifying financial challenges faced by SMEs around the world, given the new SME definition of the European Commission, effective 1/1/2005 (Appendix I). That would help the researcher develop understanding of how economic challenges might affect SMEs in various countries of the world and what might be the relationship between the stage of their national economy and the financial condition of the enterprise. In the second phase literature review was carried out in order to identify in particular, financial challenges of SMEs due to a crisis context.

According to Roman (2011) there are a series of studies that refer to the issue of financing as a barrier of SMEs. She particularly refers to the study of Beck, Demirguc-Kunt, Laeven and Maksimovic (2004), in which the authors have identified a relationship between the size and the age of an SME with its ability to access financial resources. According to their findings "young firms of small size as well as national ones face greater obstacles when they seek to obtain financial resources.", which is most probably reflecting a curvilinear relationship between size and financial chal-

lenges. Another interesting finding of the study is that "the institutional development is the most important feature that explains the differences between countries in terms of financing obstacles faced by enterprises."

Yang (2012) states that entrepreneurial firms in Sweden, South Africa, Belgium, the United States, and Norway lead the world in formal venture capital investment as a percentage of Gross Domestic Product (GDP). In contrast, China leads the world with the highest level of informal investment as a percentage of GDP. However, as it is further outlined in his study, the world financial market is not really split geographically and businesses do not enjoy custom-made solutions to finance opportunities. Even if such opportunities are accessible to larger businesses still it is very hard for SMEs who are struggling financially to design a feasible financial plan.

Okpara& Kabongo (2009), state that "According to ECA (2001), many SMEs in African countries operate in an unfriendly policy and regulatory environment, have difficulties in accessing credit, lack sufficient markets for their products, use outdated technology, lack adequate working places, and have no sufficient training facilities that are designed to meet their specific needs." This is clearly demonstrating that the blunt reality for developing countries is even uglier.

A qualitative study by Tseng and Cho, (2009) has identified that "After Lehman Brother's bankruptcy, the wave of financial crisis has been initiated (Mollenkamp, Whitehouse & Shah, 2008). The financial crisis that started in the US has become global problem and influenced the economy in both South Korea and Sweden (Fackler, 2008; The Hankyoreh, 2008). The exchange rate and export in Sweden and Korea have been significantly affected by the financial crisis. From the micro perspective, the funding cost has been increased and resulted in a liquidity problem. Many financial institutions were unable to lend money to each other due to a lack of confidence in repayments." The study of Tseng and Cho is mainly focusing on the leadership style used by leaders of financial corporations subject to economic crisis and it does not refer to SMEs whose size and value comprise a completely different profile of business. However, the findings of the study are motivating enough and provide a research interest in exploring which is the leadership style used by entrepreneurs, and owners of SMEs in a crisis context.

According to the findings of another qualitative study by Arham et al. (2013) and the conclusions drawn, "knowledge about leadership and its impact on organizational performance is still lacking" and it is expected and suggested that a qualitative study and perspective of the leadership style of the leaders of Malaysian SMEs would add significant value to the future development of the SMEs in the country.

The aim of this study is to explore the financial challenges faced by entrepreneurs and owners of Greek small sized SMEs, whose businesses are currently operating in the context of financial crisis. The lessons drawn from the Greek economy and the experiences of the entrepreneurs operating in the context of crisis might serve as lessons to learn for other economies and businesses of the same type.

Greece, is experiencing a severe economic crisis for the last 3 years, and a, the General Confederation of Greek Small Businesses and Traders (GSEVEE) estimates that a total of 61,200 small and medium-sized enterprises (SMEs) have shut down in 2012. According to the findings of its Small Enterprises Institute, it is estimated that 1,000 SMEs were closing down every week in 2012. The business environment has become very turbulent and unstable for the Greek SMEs and very few will indeed

survive in the future. The lessons taken from the Greek economic crisis will provide valuable knowledge and feedback to other European countries such as Italy, Spain and Portugal that are also entering the crisis zone. Many economists and financial analysts have expressed the concern that this could only be the beginning. Moreover, according to an article published by online Wall Street Journal (May, 2010), the Doing Business survey conducted by World Bank (2010) refers to Greece as "an economy that's hostile to free enterprise and private property, primed for corruption, lacking in labor and capital mobility, stifled by powerful trade unions and unlikely to grow without deep-rooted changes." SMEs operating in that context have been considerably impacted by the economic crisis.

A research of Marketing Research Communications (2013) conducted on behalf of GSEVEE (Hellenic Confederation of Professionals, Craftsmen and Merchants) on a sample of 1201 enterprises, revealed that 82% of the respondents answered positively that their enterprises deteriorated in the second half of 2012. The real situation in the 2nd half has been much worse than the prospects foreseen in the relevant survey of July 2012. According to the research findings, the reduction rate of the purchasing power of Greeks, is reflected on 81.1% of enterprises that show decrease in their turnover rates.

The statistics presented about Greece and the research already carried our around the financial challenges of SMEs, give to Greek SMEs a focus of interest, and it has particularly inspired the author of this paper to investigate personal experiences and individual strategies of entrepreneurs in coping with the crisis. This is further strengthened by the fact that most of the SMEs owners are also the managers of their business. That makes the majority of the SMEs dependent on the leadership style of each owner and entrepreneur. In order however, to be able to delve into the psychosocial dynamics of the entrepreneurs of SMEs, we need to be able to further understand the meaning that entrepreneurs have assigned to the crisis, the way that they perceive it and the impact on the financial viability of their businesses.

Methodology

This study used interpretive phenomenology and it was based on the interviews of two Greek entrepreneurs, within a hermeneutic phenomenological perspective. Hermeneutics is defined as the theory and practice of the interpretation of the meaning of texts (Rennie, 1999). Interpretative Phenomenological Analysis is an approach to qualitative research concerned with exploring and understanding the lived experience of a specified phenomenon (Smith, 200). Therefore, the focus of the research was to explore the perspectives of the challenges that these two entrepreneurs are facing at a time that the country (Greece) is undergoing an economic crisis. The overall goal was to explore the subjective experience of the participants in an open manner and to co-construct the meaning of the crisis.The reason that qualitative research was selected is because a semi-structured interview would reveal more in depth information of the issues raised by the interviewees, and might lead to more meaningful conclusions and findings such as:

1. personal motives and conditions for starting their business
2. opportunities and challenges they faced over the years

3. personal meaning of economic crisis and concrete examples of the way it is experienced in their business
4. personal attributions of challenges faced currently, copying strategies and future aspirations.

Moreover, the research available so far, around the research question, is mostly quantitative and generic enough, thus not providing adequate understanding of the pressures and forces exerted on the owners of entrepreneurs, in a crisis context. Moreover, as further noted in the beginning, the profiles of SMEs vary significantly according to their size and value, and they should be seriously taken into consideration when trying to investigate and identify relationships between factors that are affecting their business operation. The argument on size is further supported by the table provided via the Annual Report on the distribution of EU Small and Medium sized enterprises for 2010-2011 (Appendix II). Finally, Conger (1998) argues that "qualitative research is the cornerstone methodology for understanding the 'how' and 'why' of leadership as opposed to its 'what', 'where' and 'when'", and it is the more suitable type of research in co-constructing reality and meaning between the researcher and the participants.

Sampling

The sample of the research consisted of two individuals, selected, to abide by the following eligibility criteria:
1. Status: they should be both entrepreneurs with their own startups in any industry
2. Years of experience: at least 5 so that some kind of impact has been recorded due to the crisis
3. Size: they should employ up to ten people

The participants of the study could belong to any industry and SME sector currently operating in Greece.

A research question and an interview agenda were developed, and then the researcher first contacted people from a network of alumni of the educational institution in which this research was conducted. The researcher called five alumni in order to identify prospective participants, and the three out of the five did not fulfill the eligibility criteria. The researcher proceeded with the two individuals who agreed to participate in the interview and who fulfilled the eligibility criteria. The first participant is a male, entrepreneur/owner of a medium size SMEs in the industry of Information Technology, at the age of 43, and married. He has acquired both a Bachelor and Master's degree, and his business is operating for the last 7 years with 12 people employed. The second participant is again a male, entrepreneur/owner of a medium size SMEs in the industry of Paper, at the age of 33, and single. He has acquired a Bachelor degree and his business is operating for the last 5 years employing maximum 12 people.

Data collection

The researcher called each participant clearly explaining the purpose of the interview, and arranged a meeting for an individual interview. The meeting took place at the company of every participant, in the evening, and it was agreed that at the time of the interview they would not be distracted by any business associate or phone call. During the interview, the interviewee had to fill in a demographics questionnaire, in order to build rapport.

The interview guide developed for the interview is presented in the following table:

Table 1. Interview Schedule

Main Focus	Probe
What was your motive to become an entrepreneur?	Examples and historical reference of his life as an entrepreneur
Going back to the beginning of your business, could you refer to factors that helped you take the decision to startup your own business?	Factors that constitute internal vs external motivation
Which of these factors were controlled by you and which were externally driven?	Crisis driven factors
What is your understanding of the crisis?	Aspects of the crisis
Have you identified any impact of the crisis on your business? If yes on what aspects has the crisis affected your business?	Financial impact of the crisis
Is there anything else that you would like to add in our discussion?	

The participants of the study were informed that a follow-up session might have to be scheduled if there ambiguities during the transcription of the interview and if there are gaps and misunderstandings. With the permission of the participants the interviews were recorded and they were transcribed by the researcher.

Data analysis

The original interviews were reviewed a couple of times by the researcher who was keeping manually notes on the transcribed copies. Then the researcher proceeded with another form of analysis in order to complete coding. Originally coding was done per line and the researcher used spreadsheet analysis software in order to identify themes,

and group lines of the interview together that fell into the same theme. Then the researcher tried to identify common themes that could constitute clusters.

The analysis of data collected regarding the challenges faced by the owners/entrepreneurs of SMEs in a crisis context, produced four interrelated clusters (Appendix III):

1.   Vision of the entrepreneur: agency and independence
2.   Sources of Funding as a protective or risk factor
3.   The big challenge: International prejudice against Greek entrepreneurs
4.   Solution: International identity and exposure

These clusters were identified as the ones that were incorporating most of the issues raised by both interviewees and which seem to play an important role in the way that these leaders are perceiving challenges and opportunities during the time of a significant economic crisis.

Findings

The interviews revealed that both entrepreneurs envisioned to create their own businesses, and become successful in the future. In particular, one of the respondents said that (Int2, Ln 8-13):

"My dream was always to become an agent but international, not be based in a single country. In 2009 I found someone who was competent with computers and we were importing some products together and I asked him: "would you like to do that global?". And he said, do you want to do it small or big. During my studies at the college, I had a professor who used to tell us that if you are to do something on your own, do it big. So I told him "Big"." The main motive that they had in order to start their own businesses was to work on their own and maintain the freedom to do what they liked, without having to comply with any top management decisions that they did not agree with.

The major discussion of the survival and success of their businesses revolved around the issue of funding. The company of the first participant was completely based on external funding, coming mainly from subsidies and banks, whereas the second participant created his own business completely based on personal capital and private funds. This is directly related to the way that both participants perceive the crisis and the impact that the crisis had on each their businesses. The first participant has particularly explained the crisis as follows: (Int1, Ln115)

"My perception of the crisis is that it looks like a war. We are at a war under new terms, different from the type of wars that we were used at, at school with fights. We are at a peculiar type of war where with the use of violence we are losing or trying not to lose our personal belongings. …a group of countries maybe, I will not personalize it, and I do not really know who has the power, is it TROIKA, is it Germany, is the European Union, it does not really matter, the point is that someone is exercising pressure to the country's political leadership which constitutes the government of Greece and through them they exert pressure to the citizens of that country and they steal them, if a business goes bankrupt then the shareholders are losing their money

until a bank – which is a private business – goes bankrupt and instead of asking the shareholders to take the burden, the debt is paid by the tax payers of the country…"

This interpretation is clearly referring to an economic war where the powerful ones are setting up the rules of the game, and are actually deciding on the economic survival of the businesses. According to the participants, the defining role in the crisis is that of the banks, which have the power to support the companies financially. In the case of the first participant, whose company was totally dependent on funds coming from financial institutions and banks and investors, whose capital ownership was minimal, the impact of the crisis was devastating putting at stake even the survival of the company. The second participant on the other hand, is attributing the success and growth of his business in a time of deep economic crisis, to the fact that the company has not borrowed any money from the banks. In particular he says that they managed to survive the crisis because "we were not at all based on the banks. We were supporting our business with private capitals and this is how we managed to survive."

Concerning the crisis and Greece as a country, both participants have identified significant limitations in the way they do business, due to the fact that their companies were founded in Greece. Especially the first participant whose main customer is the public sector, has identified serious problems, given the fact that the public sector has changed considerably the way that they fulfill contracts and payments for the services and products provided to them. The first participant is quoting "what has changed to the worse is the time taken for decision making in the public sector". The second participant further says that the name of Greece in the international market has very negative connotations given that it implies crisis and lack of stability and credibility on the part of the businesses. In that context, the first participant claims that when he went abroad in order to attract international investors, the name of Greece raised significant reservations and limitations, finally discouraging foreign investors from any financial commitment. The way that the second respondent managed to overcome this difficulty was by actually opening up offices in other countries: "All problems end up in the crisis to the name of "Greece". This is not helpful at all so what we have done is that we have opened up offices in Malta, Bulgaria, UK so that we can present ourselves as internationals."

Finally, the international exposure and the links with the foreign markets and investors, seem to be the most important survival factors. According to the second respondent, the focus has always been to appeal to foreign markets and to link together foreign suppliers and customers of other countries. He has clearly defined the operation of his business as the mediator acting as the liaison and bringing together suppliers and customers from all around the world. The first respondent has also designed a survival strategy which highly depends on foreign investors. The business goal is to raise capital internally, in order to support the initial investment to travel abroad, and search for international customers and partnerships that will give them the ability to acquire funds from other countries, and eventually become crisis-resilient. Both participants confessed that confining their businesses on the Greek market alone, cannot guarantee success and survival in the future. Greece continues to remain an unstable environment, subject to many changes, which creates considerable complications in the strategic design and implementation.

Discussion and Conclusion

The above findings clearly suggest that Greece is not a favorable place to operate as a business, given the fact that the crisis and the downturn in the economy has caused considerable problems, mainly in the funding and the financial stand of the businesses.

Business owners have to overcome both the adverse economic conditions, survive the losses and adjust to a legal environment which is changing very frequently and which is impacting company goals, strategies implemented and profitability. Meanwhile, companies competing in the same industry in other countries operate in a much more friendly environment and they are continuously enjoying a competitive advantage due to the favorable economic conditions.

The opportunity for a Greek business to expand abroad is the main focus and the main survival strategy, but the Greek origin and location of operation creates a "hostile" environment given that the economic crisis of Greece has implications on issues referring to reliability, credibility and consistency.

Therefore, despite the fact that the business owners/entrepreneurs have started out with the vision to create a company that may also engage into international business, and which may be innovative enough to provide unique products and services, the adversities caused by the economic crisis may become barriers to success and even survival. This is especially true for those companies that are dependent on banks and financial institutions as opposed to the ones which are not dependent at all, and which manage to enjoy positive outcomes and growth, even in a crisis context.

Biases and Limitations

The biases of the researcher have to do with the fact that due to past studies and professional exposure, it was inherently assumed that the crisis has significant financial implications for the businesses, and that the major challenge that business/owners and entrepreneurs face is that of financial survival. Moreover, potential limitations of the specific researched study may be summarized as follows:

- rather small sample
- as a result not representative of all diverse characteristics of entrepreneurs falling under the eligibility criteria

A larger sample might have to be included in the study, in order to take into consideration entrepreneurs and owners from different industries.

The interview as a method of data collection might have been confining and it may have been more appropriate to use triangulation in relation to data analysis (different informants involved in the analysis of data and conceptualization of emergent categories of meaning). One approach of triangulation could be to confirm findings through various sources, to engage a number of researchers, or even to ask participants for validity checks.

Further research

The above findings provide adequate ground for the researcher to investigate other aspects, such as the leadership style of the owners/entrepreneurs and explore how defining that is in designing a successful strategy. Moreover, it is suggested that grounded theory is used to investigate the performance of these businesses across time, and eventually conceptualize a framework that will explain the process of acquiring resilience in a crisis context. Moreover, further investigation could be done based on qualitative analysis with more closed questions, and a larger sample size, that could lead to more generalized findings and that could constitute valuable lessons to learn.

References

Arham, A., Muenjohn, N. & **Boucher, C. (2013).** Leadership, Entrepreneurial Orientation and Performance: The case of SME's in Malaysia. *World Journal of Social Sciences*, Vol.3 (5), p.117-130

GSEVEE Report – MARC (2012) *The Depiction of the Economic Climate*, Available from http://www.gsevee.gr/deltiatupou/190-marc

The Greek Economy Explained. (2010). http://online.wsj.com/. Available from http://online.wsj.com/news/articles/

Okpara, J. O., & Kabongo, J. D. (2009). An Empirical Evaluation of Barriers Hindering the Growth of Small and Medium Sized Enterprises (SMEs) in a Developing Economy. *African Journal of Business & Economic Research*, 4(1), p. 7-21.

Rennie DL, (1999) *Qualitative research: a matter of hermeneutics and the sociology of knowledge*. *In*: Kopala M and Suzuki LA (eds) *Using qualitative methods in psychology*, pp 3-13. Thousand Oaks, Sage.

Roman, A. (2011). SMEs' Sector Access to Finance: An Overview. *Annals of Faculty of Economics*, University of Oradea, vol. 1(1), pages 431-437.

Smith JA, Eatough V 2006 Interpretative phenomenological analysis. *In*: Breakwell GM, Hammond S, Fife-Schaw C, Smith JA (eds) *Research methods in psychology*, pp322-341. London, Sage.

Tseng, P., & Cho, S. (2009). *Leadership in the 2008 Financial Crisis* (Unpublished doctoral dissertation). JÖNKÖPING UNIVERSIT. Available from: http://www.diva-portal.org/smash/get/diva2:223994/FULLTEXT01

Van Zyl, M., & du Plessis, Y. (2012). Exploring coping strategies of business leaders during an economic downturn. *South African Journal Of Economic And Management*

Sciences, 15 (4). Available from:
http://www.sajems.org/index.php/sajems/article/view/151/190

Yang, N. (2012). Small businesses and international entrepreneurship in the economic hard time: A global strategic perspective. *International Journal of Entrepreneurship*, 16, 113-131.

Appendix I

SME definition of the European Commission, effective 1/1/2005

THE NEW THRESHOLDS *(Art. 2)*

Enterprise category	Headcount: Annual Work Unit (AWU)	Annual turnover	or	Annual balance sheet total
Medium-sized	< 250	≤ €50 million (in 1996 € 40 million)	or	≤ €43 million (in 1996 € 27 million)
Small	< 50	≤ €10 million (in 1996 € 7 million)	or	≤ €10 million (in 1996 €5 million)
Micro	< 10	≤ €2 million (previously not defined)	or	≤ €2 million (previously not defined)

Appendix II

A Source: Annual Report on EU Small and Medium sized enterprises 2010-2011

1. SMEs in Greece — basic figures

	Number of Enterprises			Employment			Value added		
	Greece		EU27	Greece		EU27	Greece		EU27
	Number	Share	Share	Number	Share	Share	Billion €	Share	Share
Micro	719.952	96,5%	92,1%	1.447.218	57,6%	29,8%	27	35,3%	21,6%
Small	22.832	3,1%	6,6%	438.792	17,5%	20,4%	16	20,6%	18,9%
Medium-sized	2.893	0,4%	1,1%	264.427	10,5%	16,8%	12	15,8%	17,9%
SMEs	745.677	99,9%	99,8%	2.150.438	85,6%	66,9%	55	71,7%	58,4%
Large	563	0,1%	0,2%	362.055	14,4%	33,1%	22	28,3%	41,6%
Total	746.240	100,0%	100,0%	2.512.493	100,0%	100,0%	77	100,0%	100,0%

Estimates for 2010, based on 2002-2007 figures from the Structural Business Statistics Database (Eurostat). The estimates have been produced by Cambridge Econometrics. The data cover the 'business economy' which includes industry, construction, trade, and services (NACE Rev. 1.1 Sections C to I, K). The data does not cover the enterprises in agriculture, forestry, fishing or the largely non-market services such as education and health. The advantage of using Eurostat data is that the statistics from different countries have been harmonised and are comparable across countries. The disadvantage is that for some countries these data may be different from data published by national authorities.

Appendix III

Funding			
	Int 1	Ln133	bankrupt and instead of asking the shareholders to take the burden, the debt is paid by the
	Int 1	Ln261	More tolerance from the banks. In the past you had to pay your installment on time, and if
	Int 2	Ln69	Interviewee: "This is because we were not at all based on the banks. You were supporting
	Int 1	Ln102	Interviewee: "and focus on acquiring capital rather than implementing strategic plans
	Int 1	Ln167	You can not survive the 2%. This is not a problem due to entrepreneurship, this is due to the
	Int 2	Ln76	Of bad debts, that is we have debts from customers that we will never get them paid.
	Int 2	Ln70	Our business with private capitals and this is how we managed to survive." Pause
	Int 2	Ln12	are to do something on your own, do it big. So I told him "Big". However, big needs capital.
	Int1	Ln115	Interviewee: "My perception of the crisis is that it looks like a war. We are at a war under
	Int1	Ln120	them into consideration, and in this case there is a framework with rules which the most
	Int1	Ln132	the shareholders are losing their money until a bank – which is a private business – goes
	Int1	Ln40	one in few probabilities to succeed and enjoy potentially multiple financial benefits than
	Int1	Ln191	Me: "Therefore, this is translated to financial impact for you"…
	Int1	Ln157	Me: "So opportunities are accumulated for the ones that have the financial power.."
	Int1	Ln89	Interviewee: "They certainly turned out differently, much

			harder financially (pause) and
	Int1	Ln241	Interviewee: "We needed funding and investment. We have already developed a business
	Int1	Ln106	the crisis we did no focus at all on fund raising and suddenly the crisis, turned everything
	Int1	Ln98	the support of a group of investors so that we could focus on the operation of our
	Int1	Ln222	We practically had to borrow more money, and we produced new products, that we thought
International			
	Int2	Ln35	International"
	Int2	Ln55	can present ourselves as internationals.
	Int2	Ln6	Business abroad. Our family business is the classical type of traditional business in the
	Int2	Ln10	asked him : "would you like to do that global?". And he said, do you want to do it small
	Int2	Ln44	to the entire globe and just between these two people. Assuming that there is a buyer
	Int1	Ln232	application that we have, we approached foreign investors and potential customers, we took
	Int1	Ln231	produced. The second step we took was an effort to appeal to foreign markets, through an
	Int2	Ln8	international, not be based in a single country. In 2009 I found someone who was
Greek crisis			
	Int1	Ln15	played with my- fear of a potential failure, because a big issue, especially in Greece, is what
	Int1	Ln21	then this will not work for you. The Greek framework does not permit that, it is much
	Int1	Ln180	payment, because in the Greek public sector there were

			payment problems even before,
	Int1	Ln80	Time, but I could also start my own business later on if I was mature enough or the
	Int1	Ln221	Interviewee: "What we actually did is that we took a higher risk, without laying off people,
	Int1	Ln30	where I told him that I considered a certain practice to be risky, but bottom line "its your
	Int1	Ln84	Interviewee: "No, I wouldn't do it all, but now we have a crisis also and everything is
	Int1	Ln139	Interviewee: "Obviously entrepreneurship has changed. In the war, like the crisis, it is
	Int1	Ln155	opportunities, but that does not mean that there are opportunities for everyone due to the
	Int2	Ln174	Interviewee: "Due to the fact that – even before crisis – a big portion of the projects we are
	Int2	Ln53	Interviewee: "All problems end up in the crisis to the name of "Greece". This is not helpful at
	Int2	Ln62	Interviewee: "The crisis is not a situation that just started. It was created years before.
	Int1	Ln145	the war from the losers but in as far as business is concerned, opportunities exist for those
	Int2	Ln25	Interviewee: "I was just observing the gap. In example what the companies of my mother
Vision			
	Int1	Ln61	what could I do? Fire him? (he laughs) He would stay there, I would have to leave. So
	Int1	Ln90	Differently that our vision…
	Int2	Ln50	Interviewee: "Yes and the fact that I had the vision"
	Int2	Ln19	always on commission based business. That is not to import, not to have inventory and

The role of international networks in establishing a youth innovative entrepreneurial culture in post-communist countries: a comparison between Albania and Estonia

Gentjan Shaqiri [1] and Oliana Sula [2]

[1] Department of Management , Faculty of Economy,University of Tirana
gshaqiri@hotmail.com
[1] Department of Entrepreneurship, Estonian Business School
olasula@hotmail.com.ar

Abstract. Youth entrepreneurship is considered a pillar of development in European post-communist countries. Young people in Central and Eastern Europe are considered to be a generation that is not prisoner of the past and it is quite similar to the world's young generation who lives in this "global village". Students especially those enrolled in business programs acquire through education entrepreneurial skills, they are supposed to be more inclined towards innovation as they are conscious and adaptable in this changing world. The lack of entrepreneurial culture in post-communist countries causes a skeptic attitude towards entrepreneurship. International networks can facilitate and accommodate entrepreneurial process. This paper examines how these international networks can help in establishing a youth innovative entrepreneurial culture in Albanian and Estonian context. This study employs qualitative exploratory research methods combined with quantitative methods. A first stage of research is currently being developed and it deals with the challenges of developing networking business skills for Albanian and Estonian students, questionnaires are being elaborated in order to be distributed to Albanian and Estonian students. Action research through cross-border entrepreneurial virtual building teams between Albanian and Estonian students will be employed in a second stage of the research which will be focused in students' entrepreneurial readiness and the impact of international networks on entrepreneurial culture. Focus groups will be organized in order to discuss about the cultural aspect. This study is relevant, it fills an existing gap in the literature consisting international entrepreneurship and networks but also youth entrepreneurship culture in post-communist countries.

Keywords: *cross-border cooperation, entrepreneurial culture, international networks, post-communist countries, youth entrepreneurship*

1   Introduction

Bhave (1994), Morris et al (1994), Shane and Venkataraman (2000), Van der Veen and Wakkee (2004) and Baron (2008) conclude entrepreneurial process has 3 stages: (a) recognizing of the opportunity ,(b)the planning of the operation to take chances and (c) exploiting that opportunity. A second group of scholars represented by Shapero an Sokol (1982), Krueger et al (2000), Van Gelder et al (2006) have elaborated a 4 stages entrepreneurial process composed by : (1) intention to start, (2) recognize the opportunity, (3) preparing a chance and (4) exploit that opportunity. Vecchio(2003) created a model which resumed the entrepreneurial process: (1) the preparatory phase (pre-start), (2) start, (3) the venture being in action (exploration), (4) stopping doing business (exit).

Assuming the specific context of this research, Vecchio's model is more suitable, entrepreneurial culture consists the whole entrepreneurial process, but considering the specificity of youth target group in a post-communist countries in transition such as Albania and Estonia, the problem of the lack of initiative and motivation as a derivate of lack of entrepreneurial culture consists mainly the transition between the preparatory phase (pre-start) and the second phase the phase of start. Young people and especially students may want to be entrepreneur. They are *wantapreneurs* [1]but they fail to transform themselves in entrepreneurs. In the second section there is analysis based on the specialized theoretical framework that shows how international networks, entrepreneurial culture and youth entrepreneurship can be integrated in the context of post-communist countries such as Estonia and Albania. In the third section the use of qualitative exploratory research methods combined with quantitative is justified, the different stages of research are also shown.

2   Literature review

2.1 The definition of youth entrepreneurship culture in post-communist countries

After the 90' many scholars and academics such as Machonin (1998) have pointed that younger generations are the main source of challenge in Central and Western Europe, youth is often qualified as a "sub-culture" and it is a major social force. According to Jung (1995) the emerging individualism of Eastern European Youth will be the main condition of the creation of the new Europe.

---

[1]   Combines '*Wannabe*' and '*Entrepreneur'*. This word originated in the technology bubble of 2012 for the growing band of people who are trying to set up their own technology companies. Typically a *wantrepreneur* is still in the concept or planning stage and has yet to put together the final team, develop a product or receive funding

According to Ken Roberts (2008) the new generation in Central and Eastern Europe is more privileged that youth in Western Europe because young people in Central and Eastern Europe spend more time at the university, learn more foreign languages and are highly adaptable to change. This generation in transition has suffered many transition consequences such as high unemployment and it can be considered as a nomadic generation because some countries such as Albania experience a huge number of graduated young people, a huge labor demand but a little labor offer which creates subsequently a market failure. The young generation doesn't lack of potential but it lacks of orientation. Estonia is considered as an innovative economy in European Union.

Entrepreneurship has been considered as *best cure* against unemployment. Tyson (1994) considers entrepreneurship and managerial capacity building as one of the conditions of growth and stability in Eastern European Countries. Mc Naughton and Todorovic (2007) compare entrepreneurship in developed countries and in developing countries; they conclude that the quality of entrepreneurship matters.

. According to Morrison entrepreneurship is sub-culture able to influence and transform external environment. Verstratete (1999) and Fayolle (2003) consider entrepreneurship as culture that it is established between the interaction of values and the individual.

Institutionalists such as Bruton and Rubanik (1997) that have mainly worked with Russia estimate that the lack of institutions and good governance caused the lack of innovative entrepreneurship. Culturalists such as Mc Shane(2001) and Venkataraman (2004) stipulate that entrepreneurship is difficult to happen in traditionalist collectivist societies such as post- communist countries. When the transition from traditional entrepreneurship towards a more quality oriented entrepreneurship .Culture influences innovation and resource leverage of entrepreneurs concluded Tiessen (1997) in a study. Cultural values are factors in economic performance explaining more than half the cross national variance in economic growth according to Franke et al (1991).

Cultures have different attitudes towards entrepreneurship, risk-taking, consequences of failure, and perception about entrepreneurs. Even in entrepreneurial societies, if society doesn't reward, there is no incentive to create and innovate. Societal reflections toward entrepreneurship affect both an individual's decision to be an entrepreneur as well as society's perception, and acceptance or rejection, of entrepreneurship. Busenitz et al (2000) refer to the degree to which society admires entrepreneurial activities and thus implicitly influences its entrepreneurial spirit.

2.2   International Networks, entrepreneurship and innovation

According to Casson(1997) networks emerge as a coordination mechanism between hierarchy(firm) and the markets. Networks are supposed to be more cooperative and less competitive than the markets. Hierarchy is characterized from the lack of the re-

sources and it is quite small, markets are bigger than the networks but transactions costs are higher.

Tjosvod and Weicker(1993) stated that entrepreneurs with cooperative goals are more successful than those with an egoistic behavior. They are in the same approach of Schell (1983) that studied the level of encouragement of entrepreneurs by macro-entrepreneurs (entrepreneurial pools, incubators, linking entrepreneurs with the power elite)

Some other scholars such as Burt (1993) have focused their work in what are the benefits that the entrepreneur bring (or can have from the network) such as resources, skills, advice, equipment, reinsurance. According to Crane (1969) networks effect on opinion leadership and in the diffusion of innovation.

3   Proposed Methodology

The focus of this study is the transition between the pre-start phase and the start phase of the youth entrepreneurial process. The sample will be composed from Albanian and Estonian students.

A first part of the study has already started and it consists in the challenges of developing o business networking skills for Albanian and Estonian business students. The first data will be quantitative and qualitative collected respectively through Likert scale questionnaires and content analysis of open questions in questionnaires. The second stage includes experimental action research through creating online virtual cross boarder business teams between Albanian and Estonian students. The last stage of the research will be concentrated in the role of education in students' entrepreneurial readiness in an international network context and cultural issues will be discussed in focus groups.

The topic is particular so the study is qualitative and exploratory but it will contain also some empiric quantitative validation.

4   Conclusions

This research is necessary not only to fill the gaps of existing literature consisting entrepreneurship culture and mostly in the creation of entrepreneurial-friendly culture among young people through international networks. Students represent a specific young target group because they possess the theoretical knowledge about entrepreneurship, some of them may be *wantrepreneurs,* but a very few of them due to cultural factors do not transform themselves in entrepreneurs.

The aim of this research is to analyze the dynamics of students' network creation and how these networks can gave their contribution to the creation of an entrepreneur-

ial culture. This research is also helpful to youth policy makers dealing with the implementation of youth policies.

References

Baterman, M. (2000) Neoliberalism, SME development and the role of business supportive centers in the transitions economies of Central and Eastern Europe. *Small Business Economics*, 14(4), p.275-298.

Daughery ,C,W. (2004) Promoting Civic Entrepreneurship : The case of rural Slovakia. *Journal of the Community Development Society,* 35(1),p. 117-134.

Hafzi,M .(2003) Female Entrepreneurship in Albania. Bologna, University of Bologna.

 Naughton , Torodvic. (2007) The effect of culture, resources and quality of entrepreneurship on economic development: a conceptual framework. *International Journal of Entrepreneurship and Small Business,*4(4), p. 121-133.

Mc Millan , Woodruff (2002) The central role of entrepreneurs in transitions economies .*Journal of Economic perspectives,*16(3) ,p.153-170.

Pickering,P.(2006) Generating Social Capital for bridging ethnic divisions in the Balkans : Case study of two Bosnian Cities. *Ethnic and racial studies,*21 (1), p. 79-103.

Smallbone,Welter. (2001) The ineffectiveness of entrepreneurship in transition economies .*Small Business Economics*, 6(3), p. 165-185.

Weslund , Bolton. (2003) Local social capital and entrepreneurship. *Small Business Economics*., 21(2), p. 77-113.

Accounting for Innovation

Rezart Prifti[1]

[1] Department of Management, Faculty of Economy, University of Tirana,
`rezartprifti@feut.edu.al`

Abstract. This short paper aims to study the connection between accounting and innovation. It tries to explain where and how accounting principles would help to build a universal model for innovation. Based on the MIT model of innovation, it argues that the iterative process of innovation could be structured by an event recording mechanism such as accounting, as a mere historical artifact, in order to give the process an analyzing edge. The paper aims to study the main elements of the structure, the main issues of such an idea. This paper aims to lay groundwork for a more structured and robust study in order to analyze in great details such an idea. However, the idea discussed in this paper does not imply accounting for innovation as a measuring tool.

Keywords: *Innovation, toolkit, iterative process, decision-making*

1   Introduction

Since Luca Paciolo wrote his 27 page treatise, or even earlier during ancient times, not much has changed from bookkeeping fundaments. Change in accounting is like change in religion or religious institutions, it comes the hard way, takes time, and even when it happens doesn't change much, because usually it is an arbitrary rule that comes as a response to an industry issue.

During the last 6 or 7 years the debate on accounting rules has been extensive. Book keeping is still in action, heavy and important institutions such as ECB, Fed, IMF, BIS etc. and field experts are still debating how to cope with the issues that came into light during the crisis.

Nevertheless, the interest of this writing is not e mere history of recent facts and accounting issues *per se*, but the connection of a notion such as accounting with another notion critical to today knowledge economy: Innovation.

2   Building a framework

Before setting the ground rules with a definition of accounting or innovation let's make it clear where the two different processes should converge with each other in order to improve at least one of them, Innovation.

For a long time, innovation has been thought as a linear process (Benoit Godin, 2005), however the modern concept in the knowledge economy has changed, not that the knowledge economy affects the new point of view on innovation.

The liner model has three steps: Invention, Innovation and Diffusion (Jaffe, Newell, & Stavins, 2002) or to expand it more *Basic research – Applied research – Development – Production – Diffusion*. This view emphasizes the early stages of the process, pressing on systematic technological research, and understating the later phases. Based on the new paradigm of innovation it understates diffusion, or implementation and market components of innovation and new product development.

First, the linearity does not usually account for relations among different elements of innovation and treats them separately as they were operational processes. Second, it accounts for its history but not the essence of its creation and how something was created. Since from the early stages of the theory there were two different aspects of seeing innovation: the market pull and technology push. Both concepts together cover the two sided sources of innovation; however what happens next is the question.

Seeing innovation not any more as a linear process but more as an iterative one (Fitzgerald, Wankerl, & Schramm, 2011) the nexus with accounting would be to turn the linear work flow into a an easy and useful recording of events alias accounting, which will help later analyze and verify the chaotic iteration. The iteration includes balancing innovation elements and factors against each other while a synthesis of the process toward a goal is being reached.

Nevertheless, the idea is not to throw away the linearity and dive into chaos in order to wait for the innovative sparkle. On the contrary, taking the linear process and make it auxiliary to the new concept of innovation. And one way to do this is to account for innovation.

First let us treat accounting as a mere "recording and summarizing of business transactions", if going further we could say analyzing, verifying and reporting results also would be its functions (Jeremy, 1998). So, the accounting function would be limited to simply writing a historical line of many events in order to be analyzed and probably verified later on.

Another important aspect of accounting is measurement (Wolk, Dodd, & Tearney, 2004). It includes direct and indirect measuring. Direct measurement, which would be assigning a number to a process or an object, and indirect measurement, that has to be made by roundabout means.

Keeping in consideration the up stated aspects of accounting, let's put some perspective from the innovation side. As previously noted innovation was seen as a linear process that can be described as follows: Discovery -> Invention -> Development -> Product -> Market -> Profit. (Fitzgerald et al., 2011)

In reality, what happens during an innovation process is more chaotic than just a specific set of steps that seems more like a good willed plan or a simple historical writing of events than the real collision of different elements stimulated toward a final objective.

The whole idea is that innovation is not a straight line. Not every discovery or invention gets to be developed into a marketable product. Many researchers know that

there are plenty of inventions that never made it to the market. Even if the invention makes it through the market it is not a guarantee for profit and sustained growth.

The innovator bounces back and forth during this process, and its elements dynamically change simultaneously. Hence, the iteration of the process becomes more difficult and risk increases if not observed carefully.

If we assume that innovation has three main components such as: market, technology, and implementation, then the innovator has to consider the elements of market, technology and implementation simultaneously (for more see Fitzgerald 2010). During the iterative process these elements change as their variables do too. The innovator tries to reach a cohesive environment toward a previously specified objective, which is the innovative idea.

As any other process, innovation has a learning curve until we reach a saturation point. However it does not mean that the learning process ends. With innovation it gets transformed into abstraction. Iterating through innovation means continuously learning and abstracting, even if we get the final product we don't stop, but continue to improve in order to meet market expectations or create new ones.

There is no need to put some perspective on accounting theory. However, A.I.C.P.A. defies accounting as: "The art of recording, classifying and summarizing in a significant manner and in terms of money, transactions and events, which are in part, at least of financial character, and interpreting the results thereof".

For the purpose of this topic let's take from the definition the first part, which is: the recording, classifying and summarizing in a significant manner. In classic accounting the system would simply report a moment of the process, or a picture of company situation. In our case, accounting would be a chronology of events, transactions, decisions and effects, which in other words would be a chronology of iteration.

First, accounting for the innovation process would help the innovator to better manage the cycle. Accounting the learning curve would be a good documentation for decision making purposes. Iteration includes repeating a process in order to make it better, and documenting it would help to bring down the chaotic beginning of the innovation process. We document learning through innovation accounting and go further. This then would be a good basis for abstraction.

For example, if the innovator has done some iteration related with market elements of the product, documenting the process would be good source of decision-making and analysis for further changes of market variables. It will also be helpful for the innovator when market elements will be considered in comparison with implementation and technology elements of innovation. In other words, documenting 999 ways of a failed innovation, means crossing out 999 ways of not doing it and still looking for the right way. The record keeping function of accounting in innovation becomes more than that.

Second, documenting an innovative process means also protecting an innovation. The term protecting here takes a double meaning. On one hand it protects the whole process from defunct analysis and worthless repetition, and on the other hand it helps with the intellectual property, or the IP strategy. For example, SpaceX Company trying to build a specific product such as reusable space shuttles did not file for patents due to the specificity of the product. Since the only competitor for that product is a

government agency, revealing its secret innovation to a government agency would put the product in jeopardy. The company decided not to file for patents because the risk of disclosure was high. Therefore, the whole product innovation process has to be documented in another way.

The product of accounting is information (Leydesdorff, Rotolo, & Nooy, n.d.). Taking this in consideration makes accounting a strategic principle for every application. However, in innovation management it is crucial, because information is not only a highly priced commodity but also a fundamental part of the process. When the model is linear, a past piece of information is considered **past** and not necessarily relevant for further development phases. On the other hand, in the iterative model every piece of information is critical to dynamic decision-making and changing market realia. Invention, innovation, and diffusion are carried out simultaneously at the same time, adding also technology implementation and market information to the cocktail.

In conclusion, accounting for innovation would mean transforming the linearity of the process into a useful tool helping the complexity of the new paradigm of innovation, therefore better managing continuously changing innovation variables. The difference between invention and innovation is the whole context that the second has, but going safely and with less risk from the first to the second one need to record, summarize, analyze and measure.

Lastly, the arguments made here are in their early stages. Further development needs to be considered for the notion, also evidence and data to give shape to a more debatable idea.

3   Instead of conclusions

The notion is unexplored, or at least the connection or the approach that is being presented here. Hence, the lack of a sizable literature review. Beside that the paper does not provide empirical or other form of evidence because of the fact that this idea is in the early stage of its development.

The author is contemplating that a reasonable way to build a theoretical framework is experimenting on students groups that are creating innovative projects. Assessing from the beginning the idea and all the iteration to the point of commercialization. This approach has a big drawback because does not value the market part, the part of innovation process where iteration becomes crucial.

After surveying many work groups then one recognizes patterns, needs, and many problems that groups encountered. Hopefully this will give some insight toward the approach. Another drawback of the approach that is easy recognizable is that requires a long time range and great surveying capacity. Because of these drawbacks the idea is to think for a more conceptual approach based on actual literature, business experience, and innovation data which is becoming more available lately.

References

Benoit Godin. (2005). The Linear Model of Innovation : The Historical Construction of an Analytical Framework. *Recent Developments In The Economics Of Science And Innovation*, (30).

Fitzgerald, E., Wankerl, A., & Schramm, C. J. (2011). *Inside Real Innovation: How the Right Approach Can Move Ideas from R&D to Market-- and Get the Economy Moving* (p. 233). World Scientific.

Jaffe, A., Newell, R., & Stavins, R. (2002). Environmental Policy and Technological Change. *Environmental & Resource Economics*, *22*(1), 41–70.

Jeremy, F. (1998). Accounting and organizations : Realizing the richness of field research. *Journal of Management Accounting Researcs*, (10), 39.

Leydesdorff, L., Rotolo, D., & Nooy, W. De. (n.d.). Innovation as a Nonlinear Process and the Scientometric Perspective 1, (January 1963).

Wolk, H. I., Dodd, J. L., & Tearney, M. G. (2004). *Accounting Theory: Conceptual Issues in a Political and Economic Environment* (p. 721). Thomson/South-Western.

Environmental Indexes and Financial Ratios in SME

Savvas Savvidis [1] and Dimitrios Ginoglou [2]

[1] PhD. Candidate (Department of Accounting and Finance), University of Macedonia,
s.savvidis@avakas.com
[2] Professor (Department of Accounting and Finance), University of Macedonia,
ginogl@uom.gr

Abstract. To date, accounting and financial sciences have attempted to assess companies using financial ratios and other techniques that evaluate only the front office, without considering its impact to the environment. Recently, this assessment has been expanded to include new concepts, known as corporate social responsibility and social balance. These new concepts are usually viewed and studied using a multidisciplinary perspective with an aim to update the current and future value of the company. In this evolving scientific field a considerable effort has been made to objectively record and calculate the environmental impact of a company's activities by integrating these elements into a new form of economic balance sheet. The present paper seeks to examine the presence of weighted environmental indexes and explores the process by which they can be used to evaluate a company.

Keywords: *environmental indexes, neural networks, financial indexes*

1   Introduction

Until the last half of the 20th century, an organization's main goal was to ensure its profit and financial state without considering the impact of its relevant applications and/or productions to the environment. In fact, the single concern was how to find available and cheap natural resources, to extract them and use them efficiently. Few measures were taken to ensure that a company was environmentally friendly, not even measuring the footprint of the process done.

It was in the middle of the 20th century when it became evident that the extreme consumption of the environments resources, leads to irreversible environmental disruption. The reason for this disruption is the accumulation of environmental misuse from previous generations, as well as, current neglect. Organizations started to realize that it is for their benefit to protect the environment, and at least, be sure that the rate of environmental consumption is smaller or even equal to the rate the environment restores itself.

The new term born was "sustainable development" and firstly talked about that the Prime Minister of Norway, Gro Harlem Brundtland (Brundtland Report (1987)). The main issue now is how to satisfy present needs without risking or undermining next generations' capability of satisfying their needs. The solution to this problem demands new way of thinking, innovative choices and use of brand new technology.

Our Idea

At that point, the following idea came up: currently, all we can tell about an organization is what we know from its financial statements and/or other financial data, expressed in global standards forms. Regarding the environmental behavior, we have a blur picture since at the moment there are no obligatory regulations an organization has to follow. Still, some do keep up with some environmental standards, like GRI for instance.

We all sense that financial and environmental results have a strong interconnection, since if we had no action neither we would have financial transactions, nor environmental footprint. So, is it possible to find a model that connects them? We also sense that such a model if it exists, it will not be something simple, linear, but instead non linear and hard to define its parameters. Having all these in mind we decided to search a possible connection by using neural networks.

Financial Ratios

Financial statement analysis is important to boards, managers, payers, lenders, and others who make judgments about the financial health of organizations. One widely accepted method of assessing financial statements is ratio analysis, which uses data from the balance sheet and income statement to produce values that easily interpret financial situation (Garrison, Ray H., Noreen, E., Brewer P. C. (2009).

In fact, most organizations routinely evaluate their financial condition by calculating various ratios and comparing the results to those for previous periods, looking for differences that could indicate a meaningful change in financial condition. Another usage is to compare their current ratios with those so called as "market leaders", in order to examine if they are on the right developmental path.

There are different financial indicators measure different dimensions of financial performance, such as profitability and liquidity, and all of this information is needed to make an informed judgment about the financial health of an organization (Αδαμίδης, Α. (1998)).

To select financial performance dimensions, five top-selling textbooks on financial management were reviewed (Garrison, Ray H., Noreen, E., Brewer P. C. (2009), Horngren, C.T. &Olivar, S. (2009), Kaplan, R.,S Atkinson, A., Young, S. (2003) Βενιέρης, Γ.(2007), Αδαμίδης, Α. (1998)). Not surprisingly, there was substantial overlap and it was relatively straightforward to establish four preliminary dimensions of financial performance: liquidity, capital structure, activity and profitability.

In our case we selected to examine:

Liquidity with (1), (2)

$$\text{a) Current ratio} = \frac{CurrentAssets}{CurrentLiabilities} \quad (1)$$

$$\text{b) Quick ratio} = \frac{CurrentAssets - Inventory}{CurrentLiabilities} \quad (2)$$

Capital Structure with (3), (4), (5)

$$\text{a) Total Debt / Total Assets} = \frac{TotalLiabilities}{TotalAssets} \quad (3)$$

$$\text{b) Equity Financing} = \frac{FundBalance}{TotalAssets} \quad (4)$$

c) Cash Flow to Total Debt =

$$\frac{NetIncome + Depritiati\,onExpences}{TotalLiabilities} \quad (5)$$

Activity with (6), (7), (8)

$$\text{a) Total Asset Turnover} = \frac{TotalOperating\,Revenue}{TotalAssets} \quad (6)$$

$$\text{b) Fixed Asset Turnover} = \frac{TotalOperating\,Revenue}{NetFixedAssets} \quad (7)$$

$$\text{c) Current Asset Turnover} = \frac{TotalOperating\,Revenue}{CurrentAssets} \quad (8)$$

and Profitability with (9), (10)

a) Operating Margin =

$$\frac{TotalOperating\,Revenue - OperatingExpences}{TotalOperating\,Revenue} \quad (9)$$

$$\text{b) Return on Assets} = \frac{NetIncome}{TotalAssets} \quad (10)$$

Environmental Performance Indicators - EPI (Sustainability Indicators, Green Indexes)

Environmental Performance Indicators (EPI) have the potential to turn the generic concept of sustainability into action. Though there are disagreements among those from different disciplines, these disciplines and international organizations have each offered measures or indicators of how to measure the concept.

Various ways of operationalizing or measuring sustainability have been developed. During the last 10 years there has been an expansion of interest in EPI systems, both in industrialized and, albeit to a lesser extent, in developing countries. EPIs are seen as useful in a wide range of actors: international and intergovernmental bodies; national governments and government departments; economic sectors; administrators of geographic or ecological regions; communities; nongovernmental organizations; and the private sector.

EPI processes are supported and driven by the increasing need for improved quality and regularly produced information with better spatial and temporal resolution.

At the heart of the debate over different indicators are not only different disciplinary approaches but also different views of development. Some indicators reflect the ideology of globalization and urbanization that seek to define and measure progress on whether different countries or cultures agree to accept industrial technologies in their eco-systems. Other approaches, like those that start from international treaties on cultural rights of indigenous peoples to maintain traditional cultures, measure the ability of those cultures to maintain their traditions within their eco-systems at whatever level of productivity they choose.

Currently, main agencies around the world dealing with EPIs are United Nations Environmental Protection Agency (UN EPA), European Environment Agency (EEA), Environmental Permitting Regulations (EPR), Eco Management and Audit Scheme (EMAS) and other organizations.

Although no global standard is yet in action, as mentioned previously, there are standards brought up by individual organizations and companies that are able to collect, classify, analyze and report EPIs. The main two differences amongst them are the extension of information handled and the area of main interest given. For example, standard IPIECA (International Petroleum Industry Environmental Conservation Association) is greatly interested for the consequences of oil and pays less attention for other forms of energy consumption, e.g. electricity or gas.

Standards that are key players at the moment globally are:

a) AccountAbility, with three standards in action

AA1000APS (2008) AcountAbility Principles Standard

AA1000AS (2008) Assurance Standard and

AA1000SES (2005) Stakeholder Engagement Standart

b) IPIECA (International Petroleum Industry Environmental Conservation Association)

c) DEFRA (Department for Environment, Food and Rural Affairs – Govern. Depart. in UK)

d) FTSE4GOOD (Financial Times and the London Stock Exchange)

e) Deloitte and Touche Tohmatsu

f) ISO26000 (International Organization for Standardization)

g) ACCA (Association of Chartered Certified Accountants) and finally

h) GRI (Global Reporting Initiative)

From the above extended list, we chose to use in our research the GRI standard, since most organizations in Greece submit reports using it. From the list of all the EPI the standard has, we picked up two, the first dealing with energy and the second with water management:

a) Aspect Energy

EN3 Direct energy consumption by primary energy source

b) Aspect Water

EN8 Total water withdrawal by source

2   PROPOSED METHOD

An Artificial Neural Network (ANN) is a system based on the operation of biological neural networks, in other words, is an emulation of biological neural system. Artificial Neural Networks have different architectures, which consequently lead to different types of algorithms and problem solving, and are among the newest technologies in the finance's statistics toolbox. The field is highly interdisciplinary, but our approach will restrict the view to the financing perspective, in which ANN serve two important functions: a) pattern classifiers and b) nonlinear adaptive filters.

An ANN is an adaptive, nonlinear system that learns to perform a function (an input/output map) from data. Adaptive means that the system parameters are changed during operation, normally called the training phase. After the training phase the ANN parameters are fixed and the system is deployed to solve the problem at hand (the testing phase). The ANN is built with a systematic step-by-step procedure to follow some implicit internal constraint, which is commonly referred to as the learning rule. The input/output training data are fundamental in neural network technology, because they convey the necessary information to "discover" the optimal operating point. The nonlinear nature of the neural network processing elements provides the system with lots of flexibility to achieve practically any input/output function.

The performance of an ANN hinges heavily on the data. If data do not cover a significant portion of the operating conditions or if they are noisy, then ANN technology is probably not the right solution. On the other hand, if there is plenty of data and the problem is poorly understood to derive an approximate model, then neural network technology is a good choice.

The Biological Model

ANN emerged after the introduction of simplified neurons by McCulloch and Pitts in 1943 (McCulloch & Pitts, 1943). These neurons were presented as models of biologi-

cal neurons and as conceptual components for circuits that could perform computational tasks. The basic model of the neuron is founded upon the functionality of a biological neuron (figure 1a).
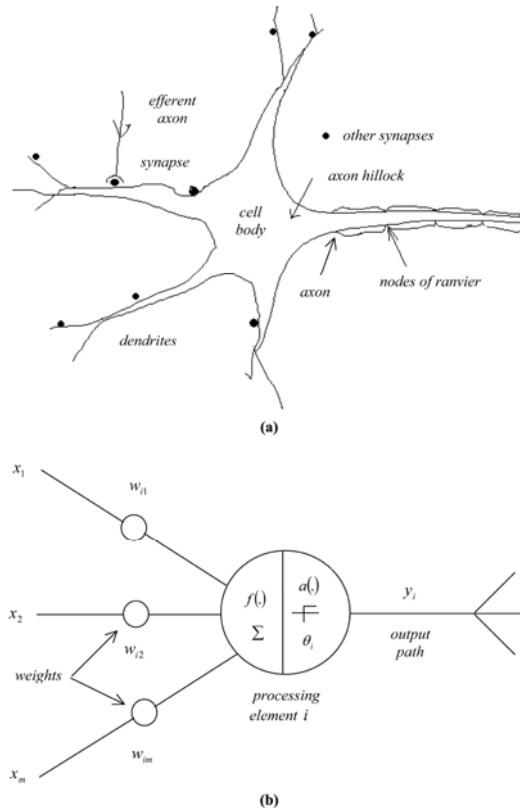
Fig. 1. Biological and Mathematical Neuron

The Mathematical Model

The functional model of the biological neuron has three basic components:

      a) Weights,

      b) A linear combination and

      c) An activation function.

   An acceptable range of output is usually between 0 and 1, or -1 and 1 and is calculated as the output of the neuron from the activation function. Mathematically, this process is described in the figure 1b, where the output of the neuron, y_i, would therefore be the outcome of activation function α (.). In our research we chose to use the sigmoid function (figure 2), since it seems it works well enough in our area of interest.
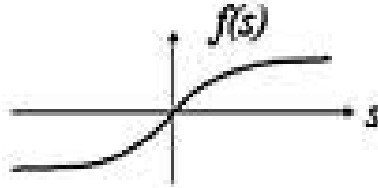
Fig 2. Sigmoid function

Supervised learning results in an adjustment of the weights of the connections between units, according to some modification rule, for example:

$$\Delta w_{jk} = \gamma y_j (d_k - y_k), \quad (11)$$

where γ is a positive constant of proportionality, representing the learning rate, d_k is the desired activation provided by data (Widrow-Hoff rule or Delta rule). Since we are using nonlinear activation functions, we use a generalised delta rule. The error measure E_p is defined as the total quadratic error for pattern p at the output units.

Back-propagation

The application of the generalised delta rule involves two phases: During the first phase the input x is presented and propagated forward through the network to compute the output values y_{po} for each output unit. This output is compared with its desired value d_o, resulting in an error signal δ_{po} for each output unit. The second phase involves a backward pass through the network during which the error signal is passed to each unit in the network and appropriate weight changes are calculated.

Backpropagation can be applied to networks with any number of layers, just as for networks with binary units it has been shown (Hornik, Stinchcombe, & White, 1989; Funahashi, 1989; Cybenko, 1989; Hartman, Keeler, & Kowalski, 1990) that only one layer of hidden units success to approximate any function with finitely many discontinuities to arbitrary precision, provided the activation functions of the hidden units are non-linear (the universal approximation theorem). In most applications a feedforward network with a single layer of hidden units is used with a sigmoid activation function for the units.

Learning rate and momentum

For practical purposes we chose a learning rate that is as large as possible without leading to oscillation. One way to avoid oscillation at large, is to make the change in weight dependent of the past weight change by adding a momentum term (Διαμαντάρας, Κ. (2007), Ζαπράνης, Α. (2005)).

3   DISCUSSION

In order to try our model, we chose to use a:
   a) Multi Layer Perceptron (MLP) model with one hidden layer and six neurons
   b) Sigmoid activation function
   c) Back Propagation training algorithm
   as seen in figure 3.
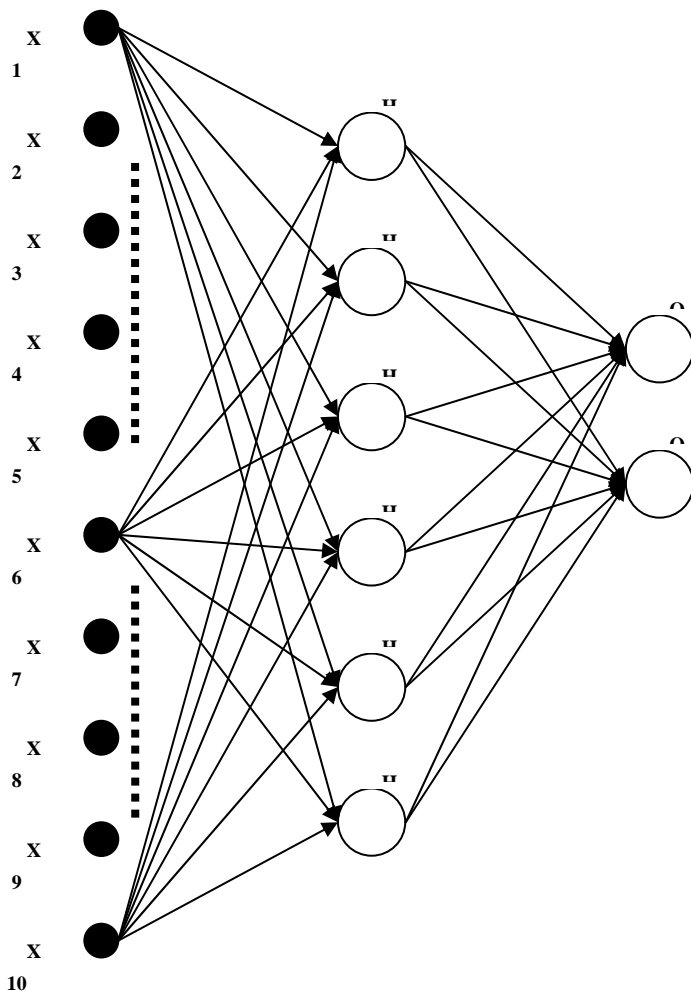
Fig 3. Out test model

If we find a realistic model that verifies out data, then we will be able to act in the reverse scheme, like an if–then–else scenario at a spreadsheet. It will allow us to decide for environmental critical projects and their extension before they are build up, explore possible limits of existing ones and in general unify the information we can have for an organization.

If on the other hand the chosen model can not verify our data, we will evaluate the error vector and decide out next step. If the error seems to be manageable, we will try different number of neurons in the hidden layer and/or more exotic training algorithm.

If finally the error seems to be unmanageable, we will change our direction to use other statistical tools, like for example data envelopment analysis (DEA).

References

Cybenko, G.V. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2, 303-314.

Funahashi, K. (1989). On the approximate realization of continuous mappings by neural networks. *Neural Networks*, 2, 183–192.

Garrison, Ray H., Noreen, E., Brewer P. C. (2009). *Managerial Accounting*. UK: McGrawHill.

Hartmann, E.J., Keeler, J.D., Kowalski, J.M. Layered (1990). Neural networks with gaussian hidden units as universal approximators. *Neural Computation, 2* 210-215.

Hebb, D. (1949). *The organization of behaviour.* UK: John Wiley.

Horngren, C.T. &Olivar, S. (2009). *Managerial Accounting*. UK: Prentice Hall.

Hornik, K Stinchcombe M & White, H. (1989). Multilayer Feedforward Networks are Universal Approximators. *Neural Networks*, 2, 359-366.

Kaplan, R.,S Atkinson, A., Young, S. (2003). *Management Accounting.* UK: Prentice Hall.

McCulloch, W.S., & Pitts, W. (1943). A logical calculus of the ideas immanent in neurons activity. *Bulletin of Mathematics and Biophysiology, 15,* 115-133.

Αδαμίδης, Α. (1998). *Ανάλυση Χρηματοοικονομικών Καταστάσεων.* Αθήνα: University Studio Press.

Βενιέρης, Γ.(2007). *Διοικητική Λογιστική.* Αθήνα: Π. Ιωαννίδου & Σια Ο.Ε

Διαμαντάρας, Κ. (2007). *Τεχνητά Νευρωνικά Δίκτυα.* Αθήνα: Κλειδάριθμος.

Ζαπράνης, Α. (2005). *Χρηματοοικονομική και Νευρωνικά Συστήματα.* Αθήνα: Κλειδάριθμος.

Global Reporting Initiative (GRI) <https://www.globalreporting.org>

ISO – Sustainability Reporting <http://www.iso.org>

AccountAbility <http://www.accountability.org>

Regional Innovation in the Sheffield City Region

Joel Capener[1]

[1]Doctoral Researcher (Sheffield University Management School), The University of Sheffield

jmcapener1@sheffield.ac.uk

Abstract. In the post-recession world innovation is increasingly championed as a key force for sustained economic recovery and growth. Owing to this recent refocusing on the importance of innovation as a growth driver, this paper introduces a study to investigate firm, regional and national dynamics that affect individual firms of differing sizes and from different sectors, ability to innovate within the context of the Sheffield City Region. This study proposes to use an inductive and qualitative approach through a series of semi-structured interviews with business owners and those engaged in innovative activities in order to understand what factors influence innovation, and thus help to create policy and organisational outcomes to support innovation-led growth. Focusing on the Sheffield City Region, this research will close the knowledge gap with regards to regional innovation.

Keywords: *Innovation, Regions, Economic Development, Firm Level*

1   Introduction

Neither the field of innovation economics, nor the interest of academics in the subject, is a new development. Joseph Schumpeter, one of the field's founding fathers wrote in 1939 that innovation involved the creation of new functions through the creation of new combinations of resources. Schumpeter also argued that the main drivers of innovation are large firms, whom possess the resources required to invest in research and development, and thus produce innovations. Today, innovation is being looked at in a new light. Instead of being seen as a process limited to large companies with formal R&D budgets and processes, research has demonstrated radical innovations can, within some industry's more than others, emerge from small businesses and entrepreneurs whom often have no official R&D framework at all (Acs and Audretsch, 2005). Perhaps most importantly in the post-recession world, innovation is being increasingly seen as a powerful, independent force on economic growth and one that is essential to create and maintain such growth sustainably into the future (Hackler, 2013). For this reason innovation as a driver of economic growth has become an increasing point of discourse and interest for academics and policy makers in recent years. This research also builds on the Regional Innovation Systems framework of Cooke (2001), which provides a basis to investigate innovation in the Sheffield City Region (SCR). Due to the complexity of this subject and its measurement this study

proposes the use qualitative methods in order to gain in depth knowledge from business and policy actors in the SCR of innovation at both a firm and regional level.

This paper will first review the existing literature on innovation, defining it based on the work of previous academic research. It will then look at the relationship between innovation and economic growth, identify the difficulty in quantifying innovation and argue the case for a qualitative approach. Next, this paper will briefly contextualize the UKs institutional and geopolitical shift away from Regional Development Agencies (RDAs) to Local Enterprise Partnerships (LEPs) which are intended to reflect functional economic areas. Finally, the literature review will define the Sheffield City Region (SCR) and assess the current state and nature of its economy in order to contextualize the proposed study. This paper will then, based on the literature review, propose the study's primary aim, objectives to be achieved and discuss the methodological approach of the study to be undertaken.

2 Literature Review

This section will first look to define the term innovation based on the work of previous academics. This is important so as to be clear what is being investigated, as well as to put the concept in a clear context. The second section will then highlight the problems of successfully quantifying innovation, and argue that a qualitative method is preferable. Finally as innovation occurs within a institutional, geographical and legislative context, this paper will examine the move away from the concept of regionalism and its rescaling towards city regions in both academia and the UK institutional context.

Defining innovation

According to Schumpeter (1939), an early innovation theorist and economist, innovation can be defined simply as the setting up through "new combinations" a new production function which can result in the production of new products, new processes, new raw materials and new forms of organisations (Lundvall, 2004). However, this definition was only adopted after Schumpeter's recognition that too narrow a definition of innovation may unfairly exclude from the academics research areas or processes that may be considered innovative. Schumpeter (1939) notes that too narrow a definition may "limit us… to the case in which the innovation consists in producing the same kind of product that had been produced before by the same kind of means of production that had been used before" (p.85). The broader definition, it is argued, takes account of other forms of innovation such as the creation of a new market, organization or commodity. Innovation as new combinations of existing resources and processes is often downplayed however, as it appears to portray innovation as an evolutionary and incremental process, thus excluding radical revolutionary innovations from its view. However revolutionary innovations often come from such recombination's of existing resources, such as with regard to the Ford Motor Company, where the four existing and well developed technologies of interchangeable parts, dedicated

machinery, the assembly line and electric motor were combined in a new formation in order to revolutionise mass production. This is a clear example amongst many of how revolutionary outcomes can come from incremental change (Hargadon, 2002).

Several academics highlight that despite varying definitions, defining a variety of innovations sub categories, a common theme emerges. It is noted that although it is a subjective term, a property which is applied to all definitions of innovation is that of "newness" (Damanpour and Wischnevsky, 2006, Garcia and Calantone, 2002; Johannessen et al, 2001). For this reason the definition used for this study will include the use of the term "new".

Hackler (2013) identifies that a distinction needs to be drawn between innovation and invention. An invention, she argues can be better defined as an idea for a new product, commodity, process or organisation, whereas an innovation can be defined as the first attempt to carry out in practice, often within a commercial context. Hacklers (2013) distinction provides a useful insight for defining innovation, in that to be considered a new innovation; it must be carried out in practice i.e. in a commercial context. Damanpour and Wischnevsky (2006), and Johannessen et al (2001) note that an innovation may be considered "new" at numerous levels, be that its initial adopter, the organisation, or the world. It may be argued therefore the degree of newness does not statically depend on the innovation itself, to be considered an innovation it must be perceived as new by the relevant unit of its adoption, be that at a firm or market level, thus giving an innovation a contextual element (Garcia and Calantone, 2002, Johannessen et al, 2001). Therefore the definition used by this study will take into account that to be considered an innovation, it must be new, carried out in practice, and therefore adopted by the relevant unit.

Finally, while sub divided in many ways, for practical purposes many studies separate an innovation as being either a new product, new process or new organisation (business) (Damanpour and Wischnevsky, 2006, Lundvall, 2004; Johannessen et al, 2001). Such an approach prevents the definition being too restrictive, while allowing academics to focus their attention on one of the three areas as innovation processes differ quite significantly depending on whether it generates a new physical product, service, process etc. Although some academics may add or rename a term, such as including a new material as a separate category (Lundvall, 2004) these three terms appear to be the most common sub divisions. "Services" while not quite as common as product, process or organisation is included as a term in some studies (Damanpour and Wischnevsky, 2006). While it may be argued that the provision of new services may be considered a new process in an organisation, or an new organisation in itself, in order to prevent its unfair omission this study will explicitly name it in the definition of innovation. Therefore taking all the above components into account the definition of innovation for this study is:

"The creation and adoption (by the relevant unit) of new products, processes, services and organisations".

Challenges of measurement

With difficulty associated with defining innovation to a suitable degree within a modern economic context comes additional difficulty in accurately measuring innovation, be that in national or regional area. Kuznets (1962) noted that a key problem that was hindering the ability of academics to suitably understand the role of technological change on the economy was the inability of scholars to measure it. Acs and Audretsch (2005) noted that measures of technological change have typically involved one of three methods, all of which have notable problems.

The first method, measuring the direct inputs into the innovative process usually in the form of R&D expenditure, is problematic as it is only a measure of input into the innovative process and not output. Expenditure on R&D provides no guarantee of an innovative output, and often such departments may have other objectives such as imitation and technology transfer. It also does not account for informal R&D which it has been pointed out by Kleinknecht (1991), is often considerable in SMEs. The second approach to measuring innovation uses an intermediate output, such as patents produced, has the problem that is noted above that not all inventions are innovations, and therefore the production of patented inventions is no guarantee of a positive economic innovative output, nor is it guaranteed that an invention which will be adopted into an innovation will be patented at all. The third method, and arguably the most accurate one, is to directly measure innovative output, often through the compilation of databases by experts looking at specific industries. While this method does not have the issues associated with the previous two methods, it is flawed in the sense that it is assumed that each innovation is assumed to be of equal importance and value to the economy. Given the problems of accurately measuring innovation, the proposed study argues that a qualitatively focused method is the most appropriate approach for looking at innovation, as this provides us with in depth data about understanding innovation and innovative activity, while avoiding the extensive methodological issues associated with measuring innovation.

From the late 1980s, based on the work of academics such as Lundvall and Nelson a body of research has been built around the concept of national systems on innovation, taking a holistic perspective focusing on interaction between different actors in the innovative process and how this interaction is influenced by a broad social, institutional and political context embedded in the national environment in which these actors operate (Teixeira, 2014). The national systems of innovation approach focuses on interactive learning between actors in the innovation process and the role that national institutions play in explaining the differences of innovation rates between different nations Asheim and Coenen (2005).

Cooke at al (1997) noted that national systems of innovation can be divided and looked at into the three parts which make up the term: National, Innovation and Systems. It is argued that within the problems exist when looking at the national component, as states may comprise of several nations and thus within the national bounda-

ries many different sub cultures and languages may exist hindering important elements of the innovation process. Thus within the regions it is "the elements of shared culture, territory and devolved administrative and/or political organisation (that) provide important dimensions of the institutional setting for innovation and other relative policy development" (Cooke at al, 1997, p.477). This therefore explains why within national systems of innovation there exists a collection of regional systems of innovation with varying degrees of effectiveness. It is therefore argued that within a state some of its characteristics can be found in certain regions, but not others within the same state. If this is the case these characteristics will also be present in the organisational characteristics of the regions firms and innovation aiding institutions, weakening the argument for national systems of innovation, and repositioning regional systems as a more appropriate unit of investigation. Asheim and Coenen (2005) argue that regions are also a particularly useful level of analysis due to the presence of globalising forces. As the world economy globalises, competitive advantage is increasingly derived from the exploitation of unique competencies, knowledge and resources, which are often tied to the regional scale. Therefore in order to effectively innovate and compete in the world economy, the regional scope of analysis is of growing relevance.

The move of Regions to City Regions

While innovation is important, within academia and the UK institutional landscape a rescaling is occurring and it is within this context that innovation needs to be understood. There is increasingly a move from the old concept of regionalism towards that of the city region as the accepted geographic framework for economic intervention. In the UK this has been reflected in the move from Regional Development Agencies built around the regional scale, and Local Enterprise Partnerships which operate on the city region scale. The following section will look at both the move of academics and institutions to the smaller city regional frame.

According to Agnew (2013), within the field of social sciences it is possible to identify five common modes of usage of the concept of "region". The concept of "new regionalism" is one of these sub divisions that has come to refers to "those theoretical and policy perspectives that directly relate the relevance of regions to the economic restructuring of national economies in the face of globalization and supranational regionalization" (Agnew, 2013, p.11).

In terms of policy implications, new regionalism is part of a body of thought that the "region" has become the main sub unit for economic development and as such needs to become a, if not the, prime focus for economic development (Webb and Collis, 2000). Although it is argued that globalisation is increasingly creating an ever homogenising and borderless world, new regionalism asserts that some flows and processes are becoming increasingly focused in a number of areas of dense economic activity (Harrison, 2007). It has been further argued that as the importance of regions as drivers of economic growth increases, the effectiveness of the nation state to act

centrally in coordinating economic policy is decreasing. This is based on the premise that the state is too large to effectively respond to rapid economic changes at a local level, and that it is those within the regions, such as business owners and policy makers, who are best placed to make decisions about the area's economic needs and act accordingly (Webb and Collis, 2000). New regionalism therefore is seen as a logical reaction to this changing economic environment through the creation of a new framework of economic governance (Agnew, 2013).

It was under this reasoning that with the election of a Labour government in 1997, efforts were made to devolve power away from the centre and to give regions more control over their affairs. One set of significant institutions that were formed as a method through which the devolution of economic strategy could be achieved were the RDAs. Pugalis (2010) States that the RDAs had five statutory purposes to:

- To further economic development and regeneration

- To promote business efficiency, investment and competitiveness

- To promote employment

- To enhance the development and application of skills relevant to employment

- To contribute to sustainable development

Most advocates of new regionalism stress the importance of gaining active engagement of the private sector in regional economic developments (Quinn, 2013), and the RDAs were designed to do this through acting as a coordinator, encouraging engagement from both public and private sector agents. It was thought that by working alongside public and business organisations a regional identity could be formed, and that it would provide an institutional capacity to shape economic process within the region it operated (Pearce and Ayres, 2007).

Despite limited successes, in the period leading up to their demise a number of criticisms have been levelled against the RDAs and their role in general. One prevalent criticism regarded the funding arrangements put in place for the RDAs by central government. Much of the funding required to deliver the regional policy goals set by Whitehall was under the control of other organisations within government, over which the RDAs were only able to exert a limited amount of influence. It is therefore argued that the amount of resources allocated to the RDAs constrained them to a degree where they were unable to make a significant impact (Pearce and Ayres, 2007; Fuller et al, 2002). Concerns were also expressed around the Geographical boundaries in which the RDAs operated, as these were based upon administrative rather than functional ones. Additionally the argument has been made that regardless of the geographical boundaries in which they operated, many RDAs perused a generic strategy,

and failed to effectively tailor their delivery to local circumstances (Huggins and Williams, 2011, Pearce and Ayres, 2007).

Quinn (2013) notes that although the definition of what constitutes a region has been widely debated amongst academics it has historically not been seen as central to debates regarding the implementation of regional economic policy. It is further noted that as the level of cohesiveness present within the regional boundaries can have a direct impact on the success or failure of policy intervention it is important to effectively define what a constitutes a functional economic area.

The formation of LEPs in 2010 under the coalition government signified an abandonment of regionalism and a move instead towards new localism, which in the view of academics such as Pugalis and Shutt (2012), served only as a mask for deep cuts in public expenditure. LEPs are in essence voluntary arrangements between various local interests such as business, civic and educational leaders whose aim is to unite disparate interests within the area in such a way that encourages growth and regeneration of their locality. They are governed by a board, which in theory represents business, civic, educational and community leaders, however in practice tend to be weighted towards either capitalist interests or local authorities. Additionally, the role that LEPs could or are meant to play remains highly ambiguous with Whitehall frequently adding to the list of functions that LEPs could perform. As a result it is a little difficult to generalise LEPs as they are at varying levels of development, pursuing different paths and strategies, have taken on different responsibilities and are using different methods to pursue their goals (Pugalis and Shutt, 2012). Through this reduction of state intervention and the reduction of regulation, it is thought that space will be created for private sector growth, and increased adaptability to local conditions (Hildreth and Bailey, 2012). While such an arrangement increases the potential flexibility of LEPs to meet the needs of their localities, whether the advantages of this outweigh the disadvantages associated with such a vague and ambiguous role is yet to be determined.

3.   Aims & Methodology

Contextualising the SCR

The Sheffield City Region is comprised of 9 local authorities: Barnsley, Bassetlaw, Bolsover, Chesterfield, Derbyshire Dales, Doncaster, North East Derbyshire, Rotherham and Sheffield (Figure 1.1). This geographical area encompasses 1.8 million people in which there are approximately 55,000 businesses, 700,000 jobs and has an estimated economic output of £24.7 Billion pounds per annum (SCR, 2013). Recent findings have demonstrated that within this area Sheffield is the key centre of employment for neighbouring towns and cities as well as the main driver for growth within the region. While other centres of employment exist such as Rotherham, Doncaster and Chesterfield these are not as strong as Sheffield, and therefore the region

can be described as weakly monocentric towards Sheffield (The Northern Way, 2009).

Currently unemployment levels within the SCR are higher than that of the UK average. In 2010, unemployment within the region as a whole was 0.7% higher than the national average. While this appears high, it should be noted that rates of unemployment vary significantly within the region, with the Derbyshire Dales having the lowest rate of 1.6% and the Dearne Valley the highest at 5.6%. Also, although the SCR experienced economic growth since 2000, compared to other northern economic centres its growth has occurred far more in the public sector than the private one (Rodgers et al, 2011). The challenge of reducing unemployment within the SCR therefore is increased by the regions high level of employment within the public sector as it is estimated that cuts in the public centre within the SCR will directly result in the loss of 14,420 jobs. This is equivalent to Sheffield loosing 2.7% of its total jobs by 2016 (SCR LEP, 2011). It is, and will continue to be therefore a major challenge in the region to work to ensure that the private sector is able to provide enough jobs to offset losses in the public sector (Rodgers et al, 2011).

While Sheffield is the main driver of economic growth in the region, with the highest levels of productivity within the region, it has less of an economic draw than other core cities. Sheffield itself is a rather self-contained city with 85% of residents both working and living in the city, taking 72% of its jobs. When compared to Manchester where 73% of residents live and work in the city taking only 31% of its jobs, this demonstrates Sheffield's comparatively weak economic draw. This is partially because although Sheffield is a relatively attractive place to work within the local area, the stronger economic centres of Manchester and Leeds are able to pull in more mobile workers as they offer higher wages (The Northern Way, 2009). As a result SCR has issues in retaining graduates than other areas, and although Sheffield has the highest concentration of graduates within the city region, it still falls below the national average.

Barnsley 1
Bassetlaw 2
Bolsover 3
Chesterfield 4
Derbyshire Dales 5
Doncaster 6
North East Derbyshire 7
Rotherham 8
Sheffield 9

Map credit: Wikipedia User AxG

Fig. 1: Map of the SCR (from: http://www.sheffieldcityregion.org.uk/about/overview)
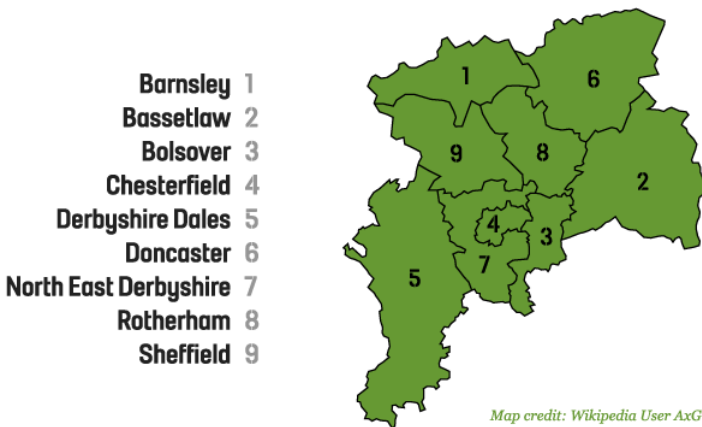
Within the SCR there a number of other important employment centres that pull in significant numbers of commuters, such as Chesterfield and Doncaster (The Northern Way, 2009). Therefore the SCR can be considered to have a fragmented economy, with a weakly monocentric pull from Sheffield (Rodgers et al, 2011). Sheffield and Rotherham have a particularly strong relationship in terms of labour links, and it has been argued that the two cities have begun to operate as one functional economic area.

In terms of its economic composition, the SCR has developed a much more knowledge based economy than previously, and while its traditional industries such as manufacturing and steel production remain prevalent, these are considerably more capital, rather than labour intensive (SCR, 2013, The Northern Way, 2009). Sheffield has developed a service sector to a higher degree than its neighbours. This has resulted in Sheffield developing business links with surrounding areas, providing accountancy business and legal services to other businesses within the region, to an extent further demonstrating Sheffield's monocentric nature. Links also exist between the regions advanced manufacturing industries, due to shorter supply chains and the build-up of trust between local firms. As a result 30% of purchases within the metal sector are made locally (The Northern Way, 2009).

The SCR, both historically and currently, has a low level of entrepreneurship and innovation relative to the rest of England. In terms of the number of businesses to its population, in the year 2000 the SCR had a gap of 29.3% compared with the national average. While there has been a general improvement in this, with the gap reducing to 27.4% below the national average in 2010, it is evident that more work needs to be done to close this gap (SCR, 2011).

Despite this, in the 2000 to 2010 period the overall increase in business stock within the SCR was 12.1% compared with the national average of 10.4%, demonstrating that gains in the region were being made and that if this is maintained the entrepreneurship gap can be closed. It is important to note that while gains are being made in the region as a whole, growth rates of business start-ups vary depending on the region. In 2010 growth rates (in terms of new starts) in North East Derbyshire were the highest in the region at 19.8% however fell by 8.5% in Bolsover. This demonstrates that when looking at enterprise a need to consider not just the SCR as a whole, but its individual parts as well (SCR, 2011).

Although Sheffield is viewed to have the potential to be an innovative economy, due to its strengths in the business service sector and advanced manufacturing as well as the University of Sheffield providing a good skills and research base, Sheffield is not believed to be reaching its full innovative potential. This is because the region still has a relatively low GVA and despite its strong advanced manufacturing sector, has a significantly lower expenditure on R&D than that of the national average (Oxford Economics, 2013). It is also identified that within the SCR Barnsley and Doncaster have characteristics of what can be described as a "low innovation city" which

demonstrates the low innovation levels of the SCR region as a whole. "low innovation city's" are areas that have a below average GVA and productivity, a low skilled work force, negative growth in the private sector and an over reliance on the public sector for the generation of jobs and growth. As described previously, although it is difficult to accurately measure innovation due to its unclear definition and issues regarding R&D expenditure as a measure of innovative output, these features still demonstrate that the SCR is not an innovative region. Given this information, the low level of business start-ups and proportion of businesses to the total population, combined with the issues relating to path dependency demonstrates that the SCR lacks the critical mass of innovative companies required to foster innovative growth.

Aims

This study's broad aim is twofold. First it aims to contribute to the body of literature regarding the innovative capacity of firms. Secondly it looks to investigate the level and nature of innovation within the Sheffield City Region. While research has been carried out on Sheffield's economy in numerous areas, to the author's knowledge no such research exists which focuses solely on Sheffield's innovation economy. Given the increasing role that innovation is being given in encouraging sustained economic growth, combined with the extended belief that the region is the most effective unit of economic analysis, makes a lack of literature on this subject a worrying deficiency and an area which warrants further investigation. This, it can be argued, is particularly true for the SCR over other areas, as traditionally the region has lagged behind the national average in a number of key economic factors, and therefore areas which have potential to increase economic growth and close this gap are particularly worthy of investigation. It is believed that the proposed research on the SCRs innovation economy will help to identify strengths that the region and firms may build upon, areas that they may improve upon, and ultimately help to generate policy and firm level outcomes to stimulate the SCRs innovation economy, and broader economic growth as a result.

As an economy is made up of many individual firms and actors, the proposed research will first begin investigating the factors that influence individual firm's ability to innovate. Use of the term "ability to innovate" allows us to take a broad and inductive approach to understanding innovation, as it allows us to explore both positive and negative factors on innovation, while leaving open the ability to explore factors which affect innovation but are not necessarily clearly one or the other. Therefore the aim of this study is to assess the factors which impact upon innovation at the firm level, and identify barriers to its generation.

In order to explore a range of factors, this aim is sub divided into four separate objectives. This will allow us to look at the factors affecting innovation from a number of different perspectives and viewpoints with regards to the firms individual, geopolitical, geo-economic and industrial context.

Objective 1.1: To identify firm specific factors which affect the individual firms' ability to innovate.

Objective 1.2: To identify region specific factors which affect the individual firms' ability to innovate.

Objective 1.3: To identify nationally specific factors which affect the individual firm's ability to innovate.

Objective 1.4: To identify industry and platform specific factors which affect the individual firm's ability to innovate.

Objective 1.5: To identify firm-size specific factors which affect the individual firms ability to innovate.

Proposed Methods

Given the problems of quantifying innovation, this study proposes the use of qualitative methods and an inductive approach. The merits of an inductive approach in this instance is that it prevents any research undertaken from being too prescriptive, potentially limiting the scope of research to prohibitive theory and unintentionally blinker attention away from potentially useful lines of enquiry, which will allow the study to gain in depth information on the subject. It is argued that the reason an in depth approach is more favourable than a quantitative measured approach is for several reasons. First, as discussed above the measurement of innovation is an area fraught with problems reducing the validity of such a method, all of which possess a deficiency in some form. This paper also argues that sufficient research has been conducted to demonstrate the importance of innovation in the stimulation of economic growth, and that therefore the measurement of innovation is less important than the study of the innovation process and how it may be encouraged in firms through the removal of barriers and capitalising on strengths. Therefore an in-depth approach is a better method through which to identify the widest range of barriers and strengths in a variety of organisational context, allowing the largest number of themes to be drawn and proposals to be generated. In summary, while an in depth approach will provide an analysis of innovation related issues on the firm level out of which policy and organisational outcomes may be drawn, measuring it will only provide a descriptive snapshot of the economy at one moment in time. Finally, in the field of innovation studies, there exists a heavy bias towards qualitative methods. This study it is believed will produce novel findings through the use of qualitative methods, readdress the imbalance of quantitative to qualitative studies currently present in the field, and produce meaningful results easily interpreted and applied by both firms and policy makers at the regional level.

The main method proposed in order to answer the objectives is to hold a series of in depth and semi structured interviews with business leaders and employees engaged

in and involved in the innovation process. The sample will include as wide a range of businesses that can be achieved given the PhDs time limit, in order to include firms of varying size, and platform. Additionally, as little can be learned about innovation from none innovating businesses, this sample will focus on businesses which already innovate. Given the intention of this research to focus in depth on such businesses, it is proposed a sample of thirty six case studies will be sufficient, divided into groups of twelve based around Sheffield three most innovative industries. The benefit of undertaking a semi structured approach to interviewing is that it keeps the research focused on the intended subject, however allows the interviewee to focus on the issues that are most important to them, as well as being able to break off in other directions that may provide useful results to the study. Once the interviews are complete, they will be transcribed and coded in order to draw out themes from the data to help answer the proposed research objectives. As Bryman and Bell (2011, p.13) note when the researcher uses inductive methods after the initial phase of theoretical reflection is undertaken, the researcher may wish to collect further data so as to test their initial conclusions, a strategy often referred to as iterative. The benefit of this approach is it adds additional validity to the researcher's additional conclusions, and allows the researcher to refine the conclusions by feeding back the generated theory and subjecting them to further scrutiny, which helps to establish under what conditions the theory does or does not hold. For this reason, so as to ensure higher validity from any conclusions drawn and to ensure a further refinement of theory, after the initial set of interviews and thematic analysis, a second series of follow up interviews will be held in order to clarify points from the previous interviews and to test the findings from their thematic coding.

3   Conclusions

Next Steps

This study will (at the time of writing), continue to read the most recent and relevant literature on innovation in order to give the study a good grounding in the latest work undertaken in the field, which is expected to last approximately 5 months. After this, research questions will be formulated, and work will be undertaken to access as many companies and individuals engaged in innovative activities as possible within the SCR while also formulating an interview guide around which to structure the interviews, which is expected to last no more than 2 months. Next the interviews will be undertaken, thematically coded, analysed and then further follow up interviews will be held. This is expected to take approximately 10 months. Finally, conclusions will be drawn from the follow up interviews and final conclusions presented. Additionally from this data an attempt will be made to distil from it appropriate recommendations for firms and policy makers in order to aid innovation within the SCR and raise its historically low levels. This is expected to take a further 6 months to complete.

References

Acs, Z and Audretsch, D. (2005) *Entrepreneurship, Innovation and Technological Change. Prepared for Foundations and Trends in Entrepreneurship.* Jenna, Max Planck Institute of Economics Group Entrepreneurship Growth and Public Policy.

Agnew, J. (2013) Arguing with regions. *Regional Studies,* 47(1), p.6-17.

Asheim, B and Coenen. (2005) Knowledge bases and regional innovation systems: Comparing Nordic Clusters. *Research Policy,* 35(8), p.1173-1190.

Bryman, A and Bell, E. (2011) *Business Research Methods.* Oxford, Oxford University Press.

Cohen, L and Levinthal, D. (1990) Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly,* 3(1), p.128-152.

Cooke, P, Uranga, M. and Etxebarria, G. (1997) Regional innovation systems: Institutional and organisational dimensions. *Research Policy,* 26(4-5), p.475-491.

Coke, P. (2001) Regional Innovation Systems, Clusters and the Knowledge Economy. *Industrial and Corporate Change,* 10(4), p.945-974.

Damanpour, F and Wischnevsky, J. (2006) Research on innovation in organizations: Distinguishing innovation-generating from innovation-adopting organizations. *Journal of Engineering and Technology Management,* 23(4), p.269-291.

Forman-Peck, J. (2013) Effectiveness and efficiency of SME and innovation policy. *Small Business Economics.* 41(1), p.55-70.

Fuller, C, Bennett, R. and Ramsden, M. (2002) The Economic Development Role of English RDAs: The Need for Greater Discretionary Power. *Regional Studies,* 36(4), p.421-428.

Garcia, R and Calantone, R. (2002) A Critical Look at Innovation Typology and Innovativeness Terminology: a Literature Review. *Journal of Product Innovation Management,* 19(2), p.110-132.

Hackler, D (2013) Unlocking the potential of innovation and entrepreneurship: The role of local policy in cities. *In:* C. Karlsson *et al.* (1st eds) *Entrepreneurial Knowledge, Technology and the Transformation of Regions.* Abington. Routledge, p.238-259.

Hargadon, A. (2002) Brokering Knowledge: Linking Learning and Innovation. *Research in Organizational Behaviour,* 24(1), p.41-85.

Harrison, J. (2007) From competitive regions to competitive city-regions: a new orthodoxy, but some all mistakes. *Journal of Economic Geography,* 7(2), p.311-332.

Hildreth, P and Bailey, D. (2012) What are the economics behind the move to LEPs? *In:* M. Ward. and S. Hardey. *(eds) Changing Gear – Is Localism the New Regionalism?.* London, The Smith Institute and Regional Studies Association, p.25-34.

Huggins, R and Williams, N. (2011) Enterprise and regional competitiveness: The role and progression of policy. *Entrepreneurship & Regional Development,* 23(9-10), p.1-26.

Kleinknecht, A. (1991) Firm Size and Innovation: Reply to Scheirer. *Small Business Economics,* 3(2), p.157-158.

Johannessen, J, Olsen, B, and Lumpkin, G. (2001) Innovation as newness: what is .new, how new, and new to whom?. *Journal of Innovation Management,* 4(1), p.20-31

Kuznets, S. (1962) Inventive Activity: Problems of Definition and Measurement. *In:* Universities-National Bureau (1st eds) *The Rate and Direction of Inventive Activity: Economic and Social Factors.* Princeton, Princeton University Press, p.19-52.

Lunvall, B. (2004) *Dynamics of Industry and Innovation: Organizations, Networks and Systems. Presented paper to the 2005 DRUID conference held at Copenhagen, June 27-29,* Copenhagen.

Northern Way (2009) *City Relationships: Economic Linkages in Northern City Regions.* http://www.sheffieldcityregion.org.uk/research. (accessed November 1, 2013)

Oxford Economics (2013) *Sheffield City Region Baseline Report.* http://sheffieldcityregion.org.uk/wp-content/uploads/2013/10/Baseline-Report.pdf. (accessed November 1, 2013)

Parr, J. (2005) Perspectives on the City-Region. *Regional Studies,* 39(5), p.555-566.

Pearce, G and Ayres, S. (2009) Governance in the English Regions: The Role of the Regional Development Agencies. *Regional Studies,* 46(3), p.537-557.

Pugalis, L. (2010) Looking Back in Order to Move Forward: The Politics of Evolving Sub-National Economic Policy Architecture. *Local Economy,* 25(5-6) p.397-405.

Pugalis, L and Shutt, J. (2012) After regions: What next for local enterprise partnerships?. *Regions,* 286(2), p.23-25.

Quinn, M. (2013) New Labours regional experiment: Lessons from the East Midlands. *Local Economy,* 28(7-8), p.738-751.

Rodgers, P., Vorley, T. and Williams, N. (2011) Creating Economic Resilience: Harnessing Entrepreneurship, Work and Enterprise in the Sheffield City Region, *Centre for Regional Economic and Enterprise Development.*

Schumpeter, J. (1939) *Business Cycles: A Theoretical, historical and Statistical Analysis of the Capitalist Process.* London, McGraw-Hill Book Company.

SCR, (2011) *Yorkshire Forward and Economic Development Learning Legacy.* http://www.sheffieldcityregion.org.uk/wp-content/themes/lepress/Downloads/0verview_-_final.pdf. (accessed November 1, 2013)

SCR. (2013) http://www.sheffieldcityregion.org.uk/about/overview/. (accessed November 1, 2013)

Teixeira, A. (2014) Evolution, roots and influence of the literature on National Systems of Innovation: a bibliometric account. *Cambridge Journal of Economics,* 38(1), p.181-214.

Northern Way. (2009) *City Relationships: Economic Linkages in Northern City Regions.* http://www.sheffieldcityregion.org.uk/research. (accessed November 1, 2013)

Webb, D and Collis, C. (2000) Regional Development Agencies and the "New Regionalism" in England. *Regional Studies,* 34(9) p.857-864.

The Development of Cities in the Global Economy

Agron Ibrahimi[1]

[1] Department of Entrepreneurial Economy, Faculty of International Economy, Finance and
Business, University of Donja Gorica, Podgorica, Montenegro
igoni@t-com.me

Abstract. The moment of dynamic variability in the field of global economy,
emerging markets and cities which promote them, is becoming an important
factor of global economic developments. In nowadays business world, cities are
meeting a number of challenges like fast and permanent changes, competitions,
globalization, global strategic crises etc. On occasion of emerging strategy at
global level, the question is: Why is it important to think regionally? Why do
we deal with economy at global level? How do global trends affect the local
economic development? In what way should the competition be realized and
maintained? The development of enterprise through clusters, the specialization
of specific economical branches and regional identity of the city are the basic
factors for a successful economic development of cities. The main goal of this
work is to show that through theoretical aspect cities are the basic movement of
economic development of every country. The theory of economic development
of the cities has the intention of achieving and creating conditions for economic
development basically ensuring economic growth and to increase well-being of
every citizen.

Keywords: *economic development of cities, global economy, clus-
ters, regional identity*

1   The development of cities in the global economy

In the post-industrial time, cities are centers of the global economy with key words
'network' and 'flow': capital, information, goods and services, people and ideas. In a
book written by Laundry (2000) was shown that cities create opportunities and inter-
actions that help to solve problems and improve the quality of life throughout the re-
gion. Beck (1999) states that: "The global economy is an economy that abolishes the
state border" (p. 45).

In the XXI century traditional instruments of economic development (model de-
velopment 'from above') became ineffective, and cities have became increasingly
susceptible to events in the global market. Therefore, the theory of local economic
development emerges (model development 'from below'). One of the main factors
creating such a development model (model development 'from below') in the globali-
zation of economy, has led to a decrease of importance at the national level.

In the literature there are several definitions of Local Economic Development (LED), although everyone has a common understanding of local economic development as a constant and planned process, beside the local authorities other stakeholders are involved (public sector, private sector and citizens) in order to create favorable conditions for economic growth, economic development and improving the quality of life of the population.

Swinburn et al. (2001) state that "the local economic development is a process in which local participants are trying to make the best use of local resources in order to maintain existing and create new jobs, strengthen and promote their business" (p. 1).

This process requires specific coordination efforts among the participants, but also the existence of a clear strategy and policy for local economic development. Because of their common interest public and business sectors are working together to create better conditions for economic growth and higher employment insurance. The aim is to boost competitiveness and contribute to economic development.

Famous scientists of Local Economic Development theory Blakey and Bradshaw (2002) state that:

> "Local economic development can be defined as the process by which holders of local authorities in the cities work together with partners from the public, business and other sectors in order to create better conditions for economic growth and job opportunities" (p. 48).

Through this process, they establish and maintain a dynamic entrepreneurial culture and create a new community and a business environment that is able to provide a higher quality of life for all in the community.

Therefore the developed economies have created LED model which in the center of its development paradigm puts the individual as a key holder of economic decisions at the local level. In these countries the local economic development has become the basic instrument for development of individual regions and cities.

Besides the LED in the global economy, it is also known the concept of 'creative cities'. The concept of 'creative city' is a relatively new phenomenon which emphasizes that human capital is the essence, not the infrastructure or architecture of the place. Varbanova (2006) states that "this is a dynamic concept that focuses on creativity, community development and culture as a key indicator of a dynamic, lively and pleasant town, the development of which is sustainable, and as such will remain for future generations" (p. 9). The United Nations Conference on Trade and Development (UNCTAD, 2011) creative cities define as: "…urban complex where various cultural activities are an integral part of the functioning of the urban economy and social life" (p. 12).

In her work Gligorijevic (2006) claims that the key prerequisite for a city to become creative is not easy to define. Is it creative residents, organizations, artists, groups, designers and leaders, or all together?

A famous theorist of creative economy Landry (2000) states that: "The creativity center of a city is the management, planning, economy, social inclusion, cultural and regional identity, all of which provide a strategic basis for such environment" (p. 256).

Thus, the model of the creative city is not linked so much to the physical space, in a way that promotes a certain area of its value. It is in fact supposed to see the possibility of realizing creative city in our region-as a method of initiating social attractiveness, which would underpin the idea of development, changes and overall improvement. It is evident that we should come up with a different way of the previous development. How should this road look like? What are the capabilities to produce creative cities, and the development of creative industries in the broadest sense of the soil of our country? The fact is that the capital city-is a magnet for the whole country. This fact is noticeable in every small town. It is another aggravating circumstance which greatly hinders a uniform dispersion of the population, which is an important factor to give a chance to other cities.

Our city development is faced with major problems. In addition to unplanned and uneven construction of infrastructure, inadequate urban planning program we are also facing the fact that there is no practical implementation of a strategy that would bring radical changes in the field of development. Moreover, such an architectural and urban image of the average Montenegrin towns, adds to the fact that most of the young educated population, without whom there is no creative city, do not want to continue living in their home town.

When we talk about the state of the urban environment in Montenegro, it is necessary to realize the identity of the cities which can be an initial step towards the path of creative industries. Most of the activities of creative industries can promote the strategy of raising the image of the city.

One of the possible reasons why do the production, wealth and people concentrate in cities is the evolution of the business philosophy that has taken place in parallel with the process of globalization. Knowledge economy and ideas replaced the traditional concepts. It was created the so called creative era in which we live today. Cities are becoming creative centers and places that generate economic growth, mostly due to the large number of creative people (members of the creative class) who have chosen to live and work in a modern urban environment. Creative cities are those that manage to attract and retain talented. A key factor in the above process is tolerance.

In the creative era to the fore, more than ever, is involved the creativity and ability of individuals. Technology, information and innovation as development factors are implied. However, innovations and ideas do not come from an unknown direction, but a specific source-has been created by people. From an economic point of view, human creativity is a form of capital and creative capital. There are different forms of capital: physical/material capital (raw materials), investment/financial capital, human capital (education of the workforce), social capital (relationships between people who work together).

According to a theorist Florida (2005, 2007) involved in the creative capital the most important factors that stimulate economic growth are: technology, talent and tolerance. From the aforementioned concept of '3T' special importance is given to tolerance. It is about the openness and willingness of one locality to accept its diversities. The location (country, region or town) that foster 'the climate' of openness and acceptability to people from different areas, which have the ability to generate the

value of their creativity is the place in which the economic growth and development are possible.

Successful cities are tolerant to diversity, attract talented people (human capital) and have their infrastructure of the modern era, which is necessary for a better quality of life and business development.

Richard Florida sees the key factor in the phenomenon of creativity in all aspects of business and the emergence of new creative class - made up of people who are educated and those who, regardless of the level of education, are creative. Creativity is the ability to create new forms, new standards, through the use of new solutions, and ideas integrated into goods and services. Florida (2005) therefore believes that creativity is the key to create competitive advantages of the city. Another theorist, Bono (2008) argues that the essential thing in business today has become creativity. Because everything else is made available to everyone, says he. De Bono (2008) state that: "If you only hope for survival and you wonder whether your business will be more competent than the competition, no doubt you are in an unenviable position" (p. 12).

If an organization cannot do anything to prevent its competitors to improve the only thing to be done is to upraise the creativity of employees to a superior level.

"UNCTAD report on the creative economy (Creative Economy Report) for the year 2010 indicates that the creative industries have been resilient to the global economic crisis of traditional industries and the real sector. Global exports of creative goods and services doubled more in the period from 2002 to 2008 and, only in the United States reached a value of 600 billion dollars per year. Despite the 12 percent decline in global trade in 2008, world trade of creative goods and services continued to expand at an average annual growth rate of 14 percent. Creative industries are sources of inclusive economic growth." (UNCTAD, Cited in Djeric, 2012, p. 5)

The creative economy is almost entirely developed at the city level and not at the national economy, thus the competitiveness of the city creates a creative center because of its concentrated creative class.

In his work Landry (2006) claimed that due to the main goal to encourage the economic growth in the creative cities are being stimulated different forms of creative industries that basically have individual creativity, skill and talent and which have the potential to create wealth and income through the generation and exploitation of intellectual property. The concept of creative city basically has the assumption that all individuals are creative in their own way and in their field and activities in which they work can contribute to creativity.

Essentially, the city is not static but complex and dynamic system which structure maintains the shape of the economic, social and cultural life, adapted to the geographical environment in which it exists.

Cities are social phenomena the maximum concentration of which is achieved through the harmony combined of effects of the environment and the dominant social processes. Due to the interaction of the environment and social-geographical factors, each city has a pronounced individuality consistency and compatibility of the physical - geographical, spatial and functional, demographic population.

On the other hand Ohmae (2005) in his book claims that most of the cities in the world search for their role on the global stage or at least to become part of the global economy. Which is the formula for success in the global economy? The formula for success in the global economy is not too complicated. Complicated is the need to determine the well-being of the nation and the state to create new skills to work in a global business. From the investors' point of view, the number of attractive cities is increasingly growing but to reach the top is the most important thing. It will become possible only when the city can sum up its *raison detre* in one word. What can *raison detre* offer to the global stage? What makes it different? As are recognized the global brands such as Coca Cola, Nike, IBM, cities must also be recognized, at least among investors, with their most recognizable features clearly visible to the rest of the competition.

The city should recognize the strength of the global economy as a means of revitalizing and enriching its own country. It needs to remove a cancerous tumor that threatens the economy and society (corruption in the administration, reducing state bureaucracy, removal of failed public companies).

Upon what is based the prosperity of the city development? The city must be shown as worth to be taken into account. Every city knows that it` is on the buyer's market, where it has to compete with others for investment.

The world is not a place where you can run, but it is the source of opportunity and prosperity. That's why cities need to attract many immigrants with higher education. Cities in transition must change the way of looking at other countries of the world to open up the global economy in order to encourage investors to start investing in their area as soon as possible. When the city is working with the rest of the world then there is no need to be afraid of the crisis that may hit the country in which the city is located.

The success of cities in the global economy will depend on: flexibility and a good leadership.

Flexibility means that the city has to adjust time, challenges and still offer investors something new in order that it can always stay in trend.

Managers must adopt the vision that attracts lots of global investors. Also, the management of the city must have a vision that is committed to education, because an educated workforce is an essential part of any economy. You need to understand that education is not a closed process, a start and a finish, but a continuous process that lasts a lifetime.

Small towns can survive in the global economy by specializing in one branch, engaging a professional manner and by selling the product with a high price and of course the products must be of good quality. Specialization implies the hope to keep the high price of their products. Italy is a good example of how it managed to survive in the local economy through specialization and retention capabilities to offer high prices, producing items for flexible demand. Through successful clusters and specializations city should have its own branding and identity for what it will be recognizable.

In order to develop a good brand, manufacturers must be aware of the difference in local communities, and that the advantages of their home city are unique and uniquely attractive.

Is branding of a city the basis for initiating the economic growth?

The popularity of branding countries, cities and regions is directly driven by the process of globalization to their unique differentiation and positioning. The clear definition of the image of a place, its better positioning acquaintance in the international market framework, enables us to learn more about the different destinations, cultures and people. The focus shifts from the local to the global level, from branding of products and services to the branding of cities, regions and countries. In the current process of connecting the world therefore creating a large global village, it is a challenge to brand the city, meet the world with its unique characteristics, promote the right values and at the same time allow the city to develop in the right direction. To achieve these goals it is necessary that all marketing activities be designed so as to serve the socio-economic interests of the city as a destination and the region in general. The main objectives of promoting a city and region are: the development of tourism and economy, the increase of investments, and the improvement of the image of a city as a destination.

The strategic approach of branding a city demands a longer period of time so as to become distinctive and attractive. The aim of the brand strategy is to anticipate the experience of the target group therefore creating the necessary activities in order to meet their needs. At the same time it affects the improvement of life in a city and the proportion of the population and tourists towards it. When you define the brand-as the strategy of the city, it is necessary to test the sensibility of the target group and accordingly take concrete actions, draft plans, to provide investments, joint ventures, organize cultural and sports events etc. It is important to initiate and conclude an agreement with the managers of public transport, architects, urban planners, public land, with the aim of attracting tourists and future residents. The idea is to make the city more attractive. All activities must have a major impact on the internal and external public affairs and specify the direction of economic, social and cultural development. In order that a brand as the strategy of a city to be effective, it must be integrated and aligned with marketing and communications, public policy, urban planning, economic and business development. It is important that the major parties, especially politicians understand the importance of the brand as a strategy and define the direction of the development. All the decisions that people make, while buying products or using services are partly rational and partly emotional. The importance of this relationship is certainly distinctive when deciding which city to visit, in which to live, in which to invest and in which to get education.

Among the biggest advantages of a 'branded' city are: fame, popularity, income and a better sensation of citizens. The opponents of a 'branded' city state that if the city is focusing on a few selected characteristics, those which at the time are acting uninteresting and inappropriate, then those characteristics can be lost forever.

The successful growth of production in a city can be organized in the form of clusters. Through clusters economy becomes multidimensional and not one-dimensional.

The manufacturers' association in the form of clusters is a possibility to progress in the global economy.

The importance of clusters is that it allows companies to become more productive, more innovative, and more competitive than it can be when a company operates independently. The atmosphere in the cluster is very competitive, but most of the participants do not compete directly, but through horizontally networked entities of the same activities involved in different markets, that put them in a position to compete globally. So, it has to do with the expression of valuable results of complementary production, i.e. with the physical boundaries of the cluster bounded by objective economic principles rather than administrative law.

The importance of a location becomes a factor of differentiation among competitors and clusters are a good example of how a competitive advantage in the global economy can exist in the local context, in the field of knowledge, greater motivation, transactional advantages, that 'distant' competitors cannot accept therefore ca not use it as a competitive advantage. The aim of the joint enterprises in a market is not only the creation of new products, but also the realization of a value added per unit of output, which allows them to further the survival of the market. A growing number of companies around the world are trying to understand the features and benefits of the cluster and the enhancement of domestic clusters to achieve better competitiveness at the global level.

Creative clusters are a form of support for the development of entrepreneurship based on connection of the various entities operating in the field of creative industries at the local level and that creates an environment that stimulates creative activities and art. Taking into account the spatial dimension in a model of creative clusters it can be observed different forms of implementing clusters-oriented development of creative industries, depending on the social, economic and historical circumstances (e.g., cultural districts, creative locations, art areas, cultural centers, etc.). The establishment of creative clusters and strengthening of their network contributes to economic development, increasing the number of jobs, solving social problems, improving the infrastructure, tourism and rebuilding urban areas.

When it comes to the tourism industry, it is evident that in the world there are many 'natural' regions that have a range of necessary resources for the development of the tourism cluster (such as sales volume generated, employment, natural resources and alike.). However, it can only be the initial value upgraded as formal: organizational, management, innovative, productive, constructive measures of social capital (collective learning and innovation), competitive and other measures.

An outstanding example of creative development through clusters in tourism and an outstanding achievement in the region and the city under a strong leadership is the capital of the Basque Country, Bilbao. In the book by Evans (2001) who took into account the case of Bilbao city was shown that the rehabilitation process, from a strategy of development to reconstruction of implementation, was the result of cooperation and partnership of local, regional and national leadership. Bilbao has used culture as a tool or catalyst for regeneration initiatives, a phenomenon that is also well documented. In the process of urban regeneration in Bilbao, it is possible to distinguish the two phases. In the first, in the early 1990s, when the emphasis was on physical regen-

eration, brought some of the most important projects, such as cleaning the bay, promoting the port towards the sea, release the key areas of the city center, and the construction of an underground band. It was a phase in which it was founded *Bilbao Metropoli 30* (BM30), a public-private company tasked with marketing activities for that city and monitoring the progress of the strategic plan for revitalization of the metropolis of Bilbao. In 1992 was founded the company of urban regeneration, *Bilbao Ria 2000*. This was a complex partnership of major institutional bodies, from central to local government, responsible for major regeneration projects. In the second phase, from the late 1990s until the early 2000s, the strategy of regeneration in Bilbao opened the way to a stronger market orientation, with greater emphasis on resources such as branding the city and attracting world-renowned architects. According to Gonzalez (2006), beginning with the Guggenheim Museum opened in 1997, the public administration promoted major projects of urban regeneration: restoration and transformation of the former shipyard area in the city center, skyscrapers housing, a new airport and new exhibition fair-these were just some of them.

With the new Strategic Plan was made a move towards a qualitative approach to regeneration and the need to invest in knowledge and values as an effective long-term strategy promoting Bilbao as a world city at an international level. The above-mentioned process of regeneration as realized by the BM30 was based on the relationship between society knowledge and the process of globalization. The process focused on competitive foreign economic policy with strict respect for the cultural identity of the Basque Country.

Therefore, many people will find that the above-mentioned project can be executed better, in the data published by Plaza (2008) show that it was quite successful with regard to growth in tourism and the creation of new tourist image of the city:

- Guggenheim Museum has played a key role in the regeneration process; from the period 1997, when it was opened, up to 2006 the number of tourists increased from 259,234 to 1,008,774;
- The number of tourists in the province of Biscay, which is the capital city of Bilbao, has risen from about 8,000 in the 1980s to nearly 90,000 guests; For example in 2006 were realized 740 905 nights;
- In the period from 1995 to 2005 in the hotel and catering sector in metropolitan Bilbao were nearly 4,000 new employees, partly thanks to the Guggenheim Museum;

The most important thing for the development of the city is the motivation-the desire and hunger for success.

The population of each city in transition needs to know that the solution to the problem lies within the city. To change their appearance on the global stage is natural and necessary.

Global trends in a local development are: economic growth, education, development of IT infrastructure, quality of life, business climate, generating knowledge.

The connection with the above global trends can indicate very much to the creation of leadership of the city with motivation for success through the use of comparative priorities of the city, natural and human resources, creating an atmosphere for attract-

ing investments from the world through innovation. The main role of interacting with the global trends also plays the economy of the state.

Less visible or invisible threads that shrouded the development of towns are as follows: forming a dedicated space for the development of economic activities , strengthening the competitiveness of enterprises and entrepreneurs, attracting capital in the local area, strengthening of local entrepreneurship, creating new jobs, designing professional personnel support, increasing source of income municipalities, businesses and personal income of employees, raising the level of education of the local population, and the possibility of further economic and spatial development of the city, and the possibility of rapid investment.

2   Conclusion

Based on the factors that promote economic growth and development of cities, we can draw conclusions about their contribution to the development of cities in the global economy.

Nowadays economic models do not operate in isolation and therefore it is necessary to create new approaches of development that take into account their identity, economic aspirations of social differences and inequalities as well as technological development or underdevelopment. The concept of development on the model of the LED and the 'creative cities' is one of the ways for future development of cities in the global economy.

Thus, creativity is still a key factor for the development of competitiveness and creative sector is becoming very important for the sustainability and continuous growth of any economy, basically the local economy. They are introducing new ways of thinking and acting in which the primary input needs individual talent or skill. These inputs can be known as unusual; but it has been observed that the rapid growing development of paradigm links the economy and culture, and combines economic, cultural, technological and social aspects of development at both the macro and micro level.

The way of connecting businesses through clusters, contributes not only to the harmonious regional development, but it is the key to the overall economic development of cities. Through specialization industrial towns can succeed to the top and create a good regional identity that attracts investors for economic growth, economic development and welfare of the population.

References

Beck, U. (1999) *What is Globalization?*, Polity Press, Cambridge.

Blakely, J.E. and Bradshaw K.T. (2002) *Planning Local Economic Development, Theory and Practice*, SAGA Publication, Thousand Oaks.

Bono, D.E. (2008) Lateral Thinking. *Profit*, Wednesday October 15 2008, p. 10.

Djeric, A. (2012) Creative Industries and the Competitiveness of Cities, *Academica-Research and Creative Economy*, Beograd.

Evans, G. (2001) *Cultural Planning: An Urban Renaissance*, Routledge, London.

Florida, R. (2005) *Cities and Creative Class*, Routledge, New York.

Florida R. (2007) *The Flight of the Creative Class*, Harper Collins Publishers, New York.

Gligorijevic, Z. (2007) Forces and Trends Shaping the Contemporary City: The Creative Sector in Creative Cities, in: Svob-Djokic, N. *Cultural Transitions in Southeastern Europe. The Creative City: Crossing Vision and New Realities in the Region. Proceeding of the 2006 International Conference held at the Dubrovnik.* Zagreb, Institute for International Relations, pp. 19-38.

Gonzalez, S. (2006) Scalar Narratives in Bilbao: A Cultural Politics of Scales Approach to the Study of Urban policy. *International Journal of Urban and Regional Research*, 30(4), p. 836-57,

Landry, C. (2000) *The Creative City: A Toolkit for Urban Innovations*, Comedia, Earthscan publications, London.

Landry, C. (2006) Lineages of the Creative City, *Research Journal for Creative Cities* (RJCC), 1(1), pp. 15-23.

Ohmae, K. (2005) *The Next Global Stage: Challenges and Opportunities in Our Borderless World*, Pearson Education, New Jersey.

Plaza, B. (2008) On Some Challenges and Conditions for the Guggenheim Museum Bilbao to be an Effective Economic Re-activator, *International Journal of Urban and Regional Research*, 32(2), pp. 506-515.

Swinburn, G., Goga, S. and Murphy F. (2001) *Local Economic Development: A Primer Developing and Implementing Local Economic Development Strategies and Action Plans,* The World Bank, Washington DC.

Varbanova, L. (2007) Our Creative Cities Online. in: Svob-Djokic N. *Cultural Transitions in South-eastern Europe. The Creative City: Crossing Vision and New Realities in the Region. Proceeding of the 2006 International Conference held at the Dubrovnik.* Zagreb, Institute for International Relations, pp. 9-18.

UNCTAD (2011) *Creative economy Report 2010: A Feasible Development Option.* [online]. Available from: http://unctad.org/en/docs/ditctab20103_en.pdf. [Accessed 15th January 2014].

Economic Integration Theories and the Developing Counties[1]

Eduard Marinov[1]

[1]International Economics Department, Economic Research Institute at BAS
eddie.marinov@gmail.com

Abstract. Economic integration theory goes through two development stages each of which addresses the relevant for its time political and economic context The first stage is regarded as classic theory or static analysis and includes the traditional theories of economic integration that explain the possible benefits of integration. The second stage includes the new economic integration theories that are often referred to as dynamic analysis of economic arrangements. Besides these two, there is a third type of integration theories that deals with the effects, benefits and constrains of economic integration arrangements of developing and least developed countries. The current paper tries to come up with a conclusion on what parts of the classic and new integration theories are applicable to integration arrangement among developing countries and to summarize these theories.

Keywords: *Economic Integration, Integration Theory, Developing Countries Integration*

1   Introduction

Regional economic integration is one of the main trends in the development of international economic relations in the last few decades. There are multiple examples, practically everywhere in the world, which demonstrate that it is not an isolated event, but an actual global phenomenon. The opportunities that are presented by the different forms of economic integration arrangements are growing and so are the means and ways for their utilization.

There is a clear distinction between the integration processes among developed countries where mainly the classic static and dynamic effects described by classic and new integration theory are sought, and those among developing and least developed countries – where the reasoning, the expected benefits and the clear constrains to the participation in integration arrangements are different.

Being a part of the theoretical chapter of my PhD thesis "Regional economic integration in Africa" the current paper is in its essence a survey of the literature on economic integration theories. As such it does not have the ambition to be a complete research with historical data or to provide some concrete examples that the theory applied to developing countries seeks to determine, neither to present a to have a policy-oriented discussion.

First, the paper explores the relevance of theories of static and dynamic analysis in the case of developing countries. It then discusses other economic integration theories that are adjusted to developing countries and are focused more on issues related to the effects, benefits and constraints of economic integration.

The main contribution of the current paper is the systematisation of economic theories that could be applied to integration efforts among developing countries.

2 Classic and new theories on economic integration effects

Many authors claim that economic integration theory goes through two development stages each of which addresses the political and economic issues relevant for its time. The first stage includes the traditional theories of economic integration, which explain the possible benefits of integration and are often referred to as static analysis. The second stage includes the new economic integration theories, which are developed in changed economic conditions and trade environment – they are referred to as dynamic analysis of economic arrangements.

Static analysis

Research of trade integration and the explanation of theoretical issues related to preferential trade agreements are based on the seminal book by Jakob Viner "The Customs Union Issue" (1951), which is often referred to as the first study of the benefits of economic integration that analyses them critically from an economic point of view (Catudal, 1951, p.210; Salera, 1951, p.84).

Viner's study is the first one to define specific criteria for the distinction of the pros and cons of economic integration. His so-called static analysis of economic integration distinguishes the now well-known effects of trade creation and trade diversion.

One speaks of trade creation when with signing a trade agreement between two countries trade is shifted from a higher cost producer to a lower cost producer among member-states. Trade diversion occurs when imports are shifted from a lower price producer from a third country, which is not a part of the integration agreement to a higher price producer from a member-state. This happens when a common customs tariff is applied if the integration agreement protects the higher cost supplier from a member-state.

Viner claims that trade creation increases a country's welfare while trade diversion reduces it. When speaking about the role of Customs unions on increasing economic welfare he says: "…customs union is only a partial, uncertain, and otherwise imper-

fect mean of doing what a world-wide non-discriminatory reduction of trade barriers can do more fully, more certainly, and equitably…" (Viner, 1950, c. 135). What Viner's theory practically means is that countries would have motivation to participate in integration if it would possibly bring more benefits than costs, or, in other words – when integration leads to more trade creation than trade diversion.

Many researches add on to Viner's static analysis by addressing different issues of integration effects. All of them come to the conclusion that no one-sided answer could be given to the question of whether customs unions increase global welfare or not. As Meade says, "Our main conclusion must be that it is impossible to pass judgment upon customs union in general. They may or may not be instruments for leading to a more economic use of resources. It all depends upon the particular circumstances of the case" (Meade, 1955, c. 107).

Dynamic analysis

Even back in the 60's, it becomes clear that static analysis of trade creation and trade diversion is not sufficient. Viner comes to the conclusion that an unpreferential trade policy (free trade) is a far better way to liberalise trade than a customs union, or, in other words the better allocation of resources is no longer applicable as a rationale for the creation of a customs union. Static effects analysis cannot fully assess the impact of integration on welfare, thus Bella Balassa introduces a new instrument to analyse the effects of economic integration on welfare – dynamic effects analysis – as a better means of explaining the reasons and economic rationale behind the creation of customs unions and economic integration schemes as a whole. A main thesis in international economics is that free trade on competitive markets enables production and consumption efficiency globally as well as in every single country. At first, the creation of preferential trade agreements motivated by the ideas of static effects analysis is viewed as a shift towards free trade and thus is perceived as a tool to increase real income. However, this turns out not to be true – this type of analysis does not give simple answers and principles, thus the attention should be put on the dynamic analysis of economic integration (Sheer, 1981, p.53).

Balassa (Balassa, 1962) and Cooper and Massell (Cooper and Massell, 1965) are the first researchers that introduce the concept of the dynamic effects of economic integration, which adds a new dimension to the research in this area. Balassa defines the main dynamic effects of integration: "large-scale economies, technological change, as well as the impact of integration on market structure and competition, productivity growth, risk and uncertainty, and investment activity" (Balassa, 1961, p.117). Schiff and Winters summarise the definition of the dynamic effects of economic integration as anything that affects the rate of medium and long term economic growth of the participating in the integration agreement member-states (Schiff and Winters, 1998, p.179).

So far a number of recent studies (Sheer 1981; El-Agra 1988; De Melo and Panagariya 1993; Fernandez 1997; Lawrence 1997; Burfisher, Robinson, and Thierfelder 2003; UNCTAD 2007, p.54) have referred to the static effects and developments of the theory of economic integration (Viner and developments) as "old regionalism",

while "new regionalism" is represented by dynamic effects such as increased competition, investment flows, economies of scale, technology transfer, and improved productivity" (Hosny, 2013, p.139). Some researchers call the two theories "first and second" regionalism, while others seek the difference in the time frame in which the effects apply to the economies: "Short-term static effects are related to the initial shift in the behaviour of economic actors,… while long-term restructuring effects are related to the improvement of the condition for the functioning of companies and their efficiency… and competition" (Panusheff, 2003, p. 37).

New theories of economic integration are developed together with the change in global economic conditions. Lawrence (Lawrence, 1997, p.18) rightly claims that the driving forces behind   previous integration efforts (simple trade creation and trade diversion) are drastically different from the factors that stand behind recent integration development, such as private sector participation, foreign direct investment, an increasing role of services, etc. Together with these, among the main effects and factors that dynamic analysis regards as coming from the participation in integration agreements are, as follows: economies of scale (Corden,1972; Balassa and Stoutjesdijk; etc.), economies of scope (Panusheff, 2003), investment creation and investment diversion (Baldwin, Forslid, and Haaland,1995), increase of competition (Marinov, 1999), etc.

The only obvious setback of dynamic analysis is that, unlike the static one, there is no reliable method for quantitative assessment of dynamic effects.

Dynamic analysis of the effects of economic integration comes from the characteristics of today's free economy. Because of their deeper scope dynamic effects have a larger impact on economic processes than static ones. The dynamic effects of economic integration can be summarized as follows: increase of investment expenditure, sustainable increase of demand, consolidation of production and increase of its specialization, improvement of the organization and management of production and production technology, rationalisation of territorial distribution and utilization of resources, increase of production efficiency, creation of economic growth, etc. (Marinov, 1999).

3   Integration determinants in developing countries

In most cases, theories of economic integration and its benefits – of dynamic ones, but even more of static ones, are not fully applicable to integration agreements among developing and least developed countries. Meier (Meier, 1960) claims that Viner's analysis has limited or no relevance to integration among developing countries. Even Balassa (Balassa, 1965, p.16) claims that theoretical literature on economic integration issues discusses customs unions only in industrialised countries. Their problems and environment are not related to economic development, but more to relative changes of production and consumption features.

The traditional theory of economic integration relies on many factors in order to reach the conclusion that net static effects determine the welfare effects of integration. Based on them, some generalisations can be made about the motivation of countries

to participate in integration processes. This part of the paper will try to distinguish those factors and effects of economic integration agreements that are relevant to developing countries. The economic determinants of integration agreements that influence the motivation of developing countries to participate in integration, in terms of both expected gains and feared negative consequences, are presented here in three main groups – general economic, market-related and trade-related factors and effects.

General economic determinants

<u>Development perspective</u>

Many researchers claim that when it comes to developing countries, economic integration should be regarded an instrument for their economic development, and not that much as customs or even trade policy (Abdel Jaber, 1971; Balassa and Stoutjesdijk, 1975). Integration theory is more focused on better resource allocation while development theory and policy deals more with the benefits from faster economic growth in the long term and the utilization of under- or not at all employed resources and production factors. Thus in many developing countries integration efforts are aimed at or more focused on the implementation of common projects in the field of development – poverty reduction, support for the development of healthcare and education systems, achievement and preservation of regional security.

<u>Macroeconomic policy coordination</u>

Shams (Shams, 2003, p.9-10) claims that even if all trade prerequisites are fulfilled when an integration agreement among developing countries is signed, the divergence of their macroeconomic policies, combined with the lack of coordination among member-states, could reduce the potential gains of integration, especially regarding the increase of interregional trade.

The issue of macroeconomic policy coordination dates back to the studies of Kahnert (Kahnert et al, 1969) and Hirschman, (Hirschman, 1971) who argue that in order for trade agreements to be durable, participating countries should try to uniform their internal monetary and foreign exchange policies (Hirschman, 1971, p. 22) and that this could be more important in promoting trade between the member countries than the customs preferences themselves. The economic areas that should be harmonised are not only limited to macroeconomic policy, but could also include industrial, social, transport, environmental policies, etc.

<u>Size of the participating countries</u>

Traditional theory assumes that the larger (in economic terms) the participating countries are, the more substantial the benefits of integration will be. According to Abdel Jaber (Abdel Jaber, 1971, p.262) if the size of the economy is measured by the gross national product, integration benefits for developing countries are negligibly small. Balassa on the other hand claims that integration gains depend not only on the size of the countries participating in the integration arrangement, but also on their rate of economic growth. Thus, as developing economies tend to grow at higher rates than already developed ones, the benefits of integration for them would be even bigger (Balassa, 1961, p.38).

Another possible measurement of the size of the integration community is the number of population. Under this criterion, developing countries will surely benefit from integration as they are usually over populated (Hosny, 2013, p.144).

Integration effects for small countries

Kreinin claims that potential gains from economic integration can be observed more clearly in small and medium sized member-states (Kreinin, 1964, p.193-194). If integration (and trade as a whole) is carried out between a small and a large country, the benefits for the small one are bigger because there is more demand for its exports. This is very substantial when the small country is a developing one and the large country is a developed one, with higher purchasing power.

A similar view is expressed by Velko Marinov, according to who in the environment of a bigger market, the comparative advantages of small national economies are manifested in their pure nature. They improve their production and market structure and increase their efficiency. "Positive effects of the participation of small countries in economic integration are achieved in medium and long term…, they assess the positive dynamic effects as far more substantial, which justifies the short term static losses" (Marinov, 1999, p. 110-111).

The argument of the positive effect of integration on small countries has its opponents. Helleiner for instance claims that the disproportion of gains in favour of the larger country is inevitable and is a result of the disparity of the economic potential of the two countries. He argues that the small country is an unequal partner who is forced to adjust to the economic and price structure in the larger member-state (Helleiner, C.K., 1996, as in Marinov, 1999, p. 112).

Market-related determinants

The welfare effects of economic integration among developing countries should not be limited only to those on production and consumption, but should also include the potential positive impact on employment, productivity, income level, specialization, competitiveness, etc.

Employment and productivity effects

It is established that in most developing countries exists a situation of generally low productivity plus mounting unemployment (Hosny, 2013, p. 141). Therefore when there is trade diversion that leads to labour force to be transferred from low-productive sectors and activities to ones with higher value added, welfare will increase.

The integration benefits to employment are even more obvious (Sakamoto, 1969, p. 283). On one hand, they are related to the fact that the changes in the geographical distribution of production influence labour demand, and on the other hand the bigger flows of workers influences labour force supply (Longi and Nijkamp, 2007, p. 3), thus increasing employment possibilities and rates.

Production specialization

Developing countries in general are specialized in the production of primary products. According to Abdel Jaber (Abdel Jaber, 1971, p.256-257) there is nothing wrong with that as long as the economic surplus gained from this type of production could be

reallocated and invested efficiently in other sectors. That however is rarely what happens in reality, thus most developing countries adopt a trade policy of diversification and import substitution to accelerate economic growth. Balanced growth can be achieved by small developing countries by increasing the size of the market, benefiting from economies of scale, and expanding their inter-industry transactions, i.e. through economic integration. For these effects to be achieved however, a strong commitment is required – both in economic and political terms.

Protection for industrial development

According to Viner, in some cases economic integration can be seen as a step towards free trade, but in others it is one towards more protection (Viner, 1950, p.41-49). There are some researchers who argue that protection trade regimes could be beneficial to developing countries. Cooper and Massell for instance believe that the main goal of integration agreements among developing countries is to support their industrial development (Cooper and Massell, 1965, p.462). This could be achieved through protection because integration, according to Sakamoto (Sakamoto, 1969, p.283-284), is equivalent to import substitution, which is a tool to support industrial development. Cooper and Massell come to the conclusion that when assessing the effects of a customs union on each member-state, one must take into account not only the change in national income, but also the development and size of each country's industry sector (Cooper and Massell, 1965, p.468).

If two developing countries create a customs union and there is a trade diversion in industrial products, welfare from the point of view of consumption, will increase when the tariffs are removed. On the other hand – from the point of view of production – welfare will decrease (viewed as an effective use of resources) because it will be replaced with production in one of the developing member-states that is more ineffective compared to that of developed third-countries. However, if such trade diversion is combined with a common external tariff that protects domestic industry, this could lead to the development of the industrial sector in both member-states. This would be particularly useful if the two developing countries are complementary, because this way each of them will expand their industrial production to supply the market in the other one (Cooper and Massell, 1965, p.475).

According to Elkan (Elkan, 1975, p. 59 -68), however, it is likely for the benefits of integration in terms of industrial production in developing countries to be unevenly distributed among the member-states. He calls this effect "backwash" - where much of the economic benefits of integration are concentrated in one or a small number of member-states (Elkan, 1975, p. 58), while economically weaker and geographically distant countries attain less benefits compared to their partners in the community.

International competitiveness

In the past, developing countries have sought motivation for economic integration in the benefits from trade diversion and import-substituting industrialization. Later on, with the introduction of the ideas of the dynamic effects of integration, they began to find arguments for integration in the economies of scale, investment creation, technology transfer, etc. Nowadays, however, the integration initiatives of developing countries far exceed those arguments – most of them pursue policies of trade liberalization and deregulation as part of their overall stabilization programs agreed with

international organizations. This approach has the goal to make economic integration policies compatible and complementary to other policies in order to promote international competitiveness. Therefore, according to Hosni, most developing countries regard economic integration as a tool for more competitiveness in a global economy (Hosny, 2013, p. 143).

Competition and complementarity

Even Viner suggests that countries producing competing (similar) products gain more benefits from integration than those producing complementary (different) ones (Viner, 1950). This comes from the fact that the more significant the difference in the price of the same goods in the potential member-states is, the greater the benefit will be (Makower and Morton, 1961, p. 35).

This should favour developing countries, because they specialize mainly in the export of products of the primary sector, thus competing in a Viner's sense. Although this is true, the fact that the major part of their exports is directed to developed countries reduces the benefits of economic integration, because it actually does not increase the volume of intraregional trade. The very category of the products of the primary sector is too large and, if split, one can see the potential benefits of integration (Abdel Jaber, 1971, p.261). Therefore Balassa argues that Viner's understanding of the criteria for competitiveness and complementarity is not at all applicable to developing countries (Balassa, 1965, p.25). Their goal actually should be to achieve a significant degree of complementarity, thus increasing the volume of intraregional trade.

More recent studies (e.g. Inotai, 1991) continue to support the thesis that in the cases of integration between developing countries complementarity and diversity of economic structures is better. In an integration agreement between similar (competing) countries, trade comes as a result of intra-sector specialization – trade expansion of this type is observed in the developed industrial countries where the size of the market and the income rate support specialization. However, this is obviously less likely for smaller and poorer markets such as those of developing countries, and therefore integration among heterogeneous (complimentary) countries is more beneficial for them.

Trade-related determinants

Benefits of trade diversion

Many researchers argue that trade diversion could actually be beneficial to developing countries. First of all, integration increases the size of the market and helps to reduce costs through economies of scale and space. Second, import substitution assists the region as a whole to spend more foreign currency for the import of capital goods and thereby contributes to the increase of investment and economic growth (Linder, 1966; Sakamoto, 1969). Furthermore, trade diversion enables consumers to buy imported goods at lower prices after the removal of tariffs thereby increasing their savings. The effect of all these, however, must be weighed against the loss of tariff revenues (Elkan, 1975, p. 59), which is particularly important for developing

countries, since most of those countries rely on them as their main source of revenue in the budget.

Linder and Sakamoto introduce the term "effective trade diversion". According to them, if economic integration among developing countries leads to trade diversion it should not reduce welfare because the production substitution will be from an efficient developed country (outside the integration agreement) towards a relatively efficient developing member-state, thus creating benefits in terms of employment and income within the community as well.

Initial tariff levels

Meade assumes that the higher the initial rates of tariffs between countries entering an integration agreement are, the higher the expected benefits of integration among them will be (Meade, 1955) –because the removal of the tariff will have a greater impact in terms of both welfare and intraregional trade. This is specifically important when it comes to developing countries because the national tariffs of most of them are rather high, mainly due to their desire either to increase revenue or to protect national production.

International trade as share of GDP

Lipsey assumes that the lower the share of international trade in GDP of the member-states of an integration agreement is, the greater the expected benefits of a customs union on welfare will be (Lipsey, 1960, pp. 508-509). This is very important for developing countries because trade as a percentage of GDP in low-income countries has always been lower than in countries with a high level of income, although in recent years this imbalance is decreasing (Hosny, 2013, pp. 144-145). However, the same does not apply to countries with medium levels of income and least developed countries – their share of trade in GDP is even more significant than that in high-income countries. It can therefore be concluded that this criterion is not applicable to developing countries, because subgroups among them may have a larger or smaller share of trade of GDP compared with high-income countries.

Share of intra-regional trade

According to Lipsey an integration agreement will bring more benefits in terms of welfare if the share of intra-regional trade is growing, while trade with the rest of the world is decreasing (Lipsey, 1960, pp.508-509). Studies show that trade between developing countries is always much weaker than that between developed countries, suggesting that the benefits of integration regarding welfare will also be smaller.

However, other researchers (Balassa, 1965; Abdel Jaber, 1971) believe that this assumption should not always be taken for granted. They list several factors that restrict trade among developing countries, arguing that if these barriers are removed, trade flows between developing countries engaged in an integration process will likely increase. These factors include: first, the low level of economic development; second, inadequate transport infrastructure and facilities; third, foreign currency control and other restrictions on imports; fourth, inadequate marketing; fifth, the lack of standardization.

Fostering regional trade

It is widely recognized that the best indicator of the success of an integration agreement is the increase of the share of intra- and inter-regional trade in the total

trade flows of member-states. Although this is an important aspect of integration Inotai (Inotai, 1991, p.10) believes that it should not be seen as a means to its end. Equally important are the industrial development, the adequate infrastructure, the increase of the technological level, etc. Furthermore, the growth of regional trade may be the result of trade diversion from more efficient and competitive third countries. Therefore it can be regarded as positive only if it is combined with improving global competitiveness as a whole.

Change of the trade structure with developed countries

A major part of the imports from developed to developing countries consists of capital goods. From the dynamic analysis point of view, integration among developing countries requires substantial investments and since most of them are imported from developed countries in the form of capital goods it is likely that the volume of imports of integrating developing countries will grow. The conclusion of Mikesell is that the long-term goal of integration between developing countries should not be to reduce trade with the outside world, but rather to change in their trade structure (Mikesell, 1965, p.209).

Sakamoto (Sakamoto, 1969, p.293) believes that if the result of integration among developing countries is the trade diversion of consumer goods, this will release more foreign currency for imports of capital goods from third (developed) countries. The volume of trade with the rest of the world may not change or may even increase, but the important thing is it changes its structure.

Transport infrastructure

Transport costs reduce the potential benefits of trade integration across countries. Distance itself affects the inter-sectorial trade. This is particularly important for developing countries that enter into integration agreements for two reasons: first, as countries with similar income per capita are more dependent on inter-sectorial trade. Second, transport infrastructure and facilities in developing countries are often in poor condition or even missing, or, if existing, they are designed to promote the transport of export of primary sector products the from developing to developed countries.

Therefore, as Balassa argues, transport costs between two bordering developing countries may actually be higher than those between one of them and a remote developed country (Balassa, 1965, p.31). This must be taken into account when considering the integration of developing countries, thus according to Abdel Jaber (Abdel Jaber, 1971, p. 262) in the preparation of integration agreements between developing countries one should pay special attention to the issue of existing transport facilities and infrastructure.

Complex theories

Static and dynamic approach

According to many researchers, one must pay more attention to dynamic rather than to static effects when assessing integration processes among developing countries (Sakamoto, 1969; Abdel Jaber, 1971, etc.). Rueda-Junquera claims that traditional integration theory, which analyses the static effects of resource allocation, implies rather small gains for developing countries and thus the motivation for participation in

integration agreements should be sought in the dynamic analysis of integration and the effects that it reveals (Rueda-Junquera, 2006, p.3-4).

According to Abdel Jaber, traditional integration theory strongly relies on the neo-classical assumptions for full employment, perfect competition, constant returns of scale and perfect mobility of production factors (Abdel Jaber, 1971, p.264-265). Thus, the analysis is restricted just to the static effects and the dynamic ones are those that could bring a higher economic growth rate and utilization of underemployed economic potential.

The training ground theory

Some of the issues discussed above – how to increase international competitiveness, what specialization to aim at, whether to use protection to support industrial development in integration agreements between developing countries – are thoroughly discussed in the training ground theory. According to Inotai, this theory rests on the hypothesis that through the first stages of integration among developing countries their international competitiveness could be gradually improved if they depend on the regional market in the first stage of industrialization (Inotai, 1991, p.6-7). Free trade among member-states, combined with high tariffs for third countries' imports should give temporary protection to emerging industries as well as a market that is big enough to support the future industrial development. This process is referred to as "import substituting industrialization" (Rueda-Junquera, 2006, p.4) and gives enough time for the development of the industrial sectors of developing countries. The openness to global markets could be realized on a later stage when developing countries have reached a certain degree of efficiency and technological development. Therefore economic integration among developing countries could be seen as a transition stage towards an open economy and competition with the rest of the world after a short period of training, thus the theory is called "training ground" theory.

Although it looks sound from a theoretical perspective, there are some arguments against this theory. Inotai for instance (Inotai, 1991, p. 7) argues that first of all, developing regional markets in many cases are not big enough to enable industrial development in the terms of economies of scale; second, as a result of the training process there is rather small or even no improvement; third, there are great differences in demand preferences and tastes regarding the imports from third countries compared to those coming from interregional trade. Besides, there are no guarantees that developing countries would take on and fulfil the commitment to open up and liberalise their trade with the rest of the world at a certain point, thus protection measures could become permanent instead of temporary.

The package approach

Another complex way to implement integration among developing countries is the package approach. According to Balassa and Stoutjesdijk, a package approach specifically and explicitly aims at facilitating the integration process and enhancing the stability of an integration agreement by assuring that each member-state is responsible for the implementation of a single integration project within a common package of such projects (Balassa and Stoutjesdijk, 1975, p.53). These could include transport, communication, public goods, education, science, agriculture, mining, industry, etc. An important condition for the successful application of the package approach is that

comprehensive information regarding the distribution of benefits and costs of each project on each member country should be available so that there are no member-states who feel there is inequality in the gains and expenses distribution of the integration process. Balassa and Stoutjesdijk argue that although this approach may seem plausible, problems such as financing, monitoring and controlling may arise.

4   Conclusion

From the above said, it is obvious that the rationale behind economic integration among developing countries could not be defined and explained just by the static and dynamic effects that determine integration between developed economies. With developing countries some factors have a stronger, while, controversially, others have a weaker impact on their willingness to participate in integration agreements.

To assess the integration benefits and costs for developing countries one must take into account their specifics such as stage of economic development, structure of the economy, production characteristics, demand preferences, trade regimes and policies, etc.

Another thing that should be noted is that while in developed countries the main rationale for economic integration comes from economic groups of stakeholders, in developing countries integration processes often initially start as a political goal and effort, which in most cases leads to unsatisfactory economic results. The complexity of the political determinants of economic integration among developing countries and their interrelations with economic rationale will be subject to further research as will be the application of the theories presented to the real-world case of regional economic integration efforts in Africa.

References

Abdel Jaber, T. (1971) A review article: The relevance of traditional integration theory to less developed countries. *Journal of Common Market Studies* 9 (3): pp. 254-267.

Balassa, B. (1961) *The Theory of Economic Integration*. Richard D. Irwin. Homewood, Illinois.

Balassa, B. (1965) *Economic development and integration*. Centro De Estudios Monetario Latinoamericanos. Mexico.

Balassa, B. and A. Stoutjesdijk. (1975) Economic integration among developing countries. *Journal of Common Market Studies*, pp. 37-55.

Baldwin, R., R. Forslid and J. Haaland. (1995) Investment creation and investment diversion: Simulation analysis of the single market programme. *NBER Working Paper no. 5364*. Cambridge.

Catudal, H. M. (1951) Review of The customs union issue, by Jacob Viner. *The American Journal of International Law* 45 (1): pp. 209-211.

Cooper, C. A., and B. F. Massell. (1965) Toward a general theory of customs union for developing countries. *The Journal of Political Economy* 73 (5): pp. 461-476.

Corden, W. M. (1972) Economies of scale and customs union theory. *The Journal of Political Economy* 80 (3): pp. 465-475.

Elkan, P. G. (1975) Measuring the impact of economic integration among developing countries. *Journal of Common Market Studies* 14 (1): pp. 56-68.

Hirschman, A. (1971) *A bias for hope: Essays on development and Latin America*. Yale University Press. New Haven and London.

Hosny, A.S. (2013) Theories of Economic Integration: A Survey of the Economic and Political Literature. *International Journal of Economy, Management and Social Sciences* 2 (5) May 2013, pp. 133-155

Inotai, A. (1991) Regional integration among developing countries, revisited. *Policy, Research, and External Affairs Working Paper no. 643*. Washington, D.C.: World Bank.

Kahnert, F, P. Richards, E. Stoutjesdijk, and P. Thomopoulos. (1969) *Economic integration among developing countries*. Paris: Development Center of the Organization for Economic Co-operation and Development (OECD)

Kreinin, M. (1964) On the dynamic effects of a customs union. *The Journal of Political Economy* 72 (2): pp. 193-195.

Lawrence, R. Z. (1997) Preferential trading arrangements: The traditional and the new. Chapter 2 in *Regional Partners in Global Markets: Limits and Possibilities of the Euro-Med Agreements*, edited by Ahmed Galal and Bernard Hoekman. Center for Economic Policy Research and the Egyptian Center for Economic Studies.

Linder, S. B. (1966) Customs unions and economic development. In *Latin American Economic Integration*, edited by Miguel S. Wionczek. New York: Praeger.

Lipsey, R. G. (1960) The theory of customs unions: A general survey. *The Economic Journal* 70 (279): pp. 496-513.

Longi, S and P. Nijkamp. (2007) Economic Integration and Labour Markets: Ways Forward. In: *Regionalization, Growth and Economic Integration*. Physica-Verlag, pp. 3-24.

Makower, H., and G. Morton. (1953) A contribution towards a theory of customs union. *The Economic Journal* 63 (249): pp. 33-49.

Marinov, V. (1999) *Regional Economic Integration*. PH Stopanstvo, Sofia. (in Bulgarian)

Meade, J. E. (1955) *The theory of customs unions*. Amsterdam: North Holland.

Meier, G. M. (1960) Effect of a customs union on economic development. *Social and Economic Studies* (March 1960): pp. 29-36.

Mikesell, R. F. (1965) The theory of common markets as applied to regional arrangements among developing countries. *In International Trade Theory in a Developing World*, edited by Harrod and Hague. New York: St. Martin's Press.

Panusheff, E. (2003) *Economic integration in the European Union*. PH Nekst, Sofia. (in Bulgarian)

Rueda-Junquera, F. (2006) European integration model: Lessons for the Central American common market. *Jean Monnet/Robert Shuman Paper Series* 6 (4) Florida: University of Miami.

Sakamoto, J. (1969) Industrial development and integration of underdeveloped countries. *Journal of Common Market Studies* 7 (4): pp. 283-304.

Salera, V. (1951) Review of The customs union issue, by Jacob Viner. *The Journal of Political Economy* 59 (1): p.84.

Schiff, M. and L. A. Winters. (1998) Dynamics and politics in regional integration arrangements: An introduction. *The World Bank Economic Review* 12 (2): pp.177-195.

Shams, R. (2003) Regional integration in developing countries: Some lessons based on case studies. *HWWA Discussion Paper* no. 251. Hamburg, Germany: Hamburg Institute of International Economics.

Sheer, A. (1981) A survey of the political economy of customs unions. *Law and Contemporary Problem*s 44 (3): pp.33-53.

Viner, J. (1950) *The Customs Union Issue*. New York: Carnegie Endowment for International Peace.

Foreign Direct Investment and its impact on Economic Growth in Southeastern European Economies, empirical investigation

Kushtrim Zeqiri

[1] City College (Regional Town and Development), University of Sheffield
zkustrim@seerc.org

Abstract. The following paper will discuss the much debated topic regarding FDI and economic growth. The following paper is an empirical investigation, with regards to Southeastern European economies (Croatia, Serbia, Bosnia and Herzegovina, Montenegro, Albania, Kosovo, and FYROM-Macedonia). The empirical investigation will cover the period from year 2000 to 2013. This study aims to observe the relationship between FDI and economic growth. There are three steps applied for examining the relationship among the variables. First step, tests for examining the order of integration among variables. Second step, examine the long run relationship of the variables. Final step, testing for short-run causality.

Keywords: *Foreign Direct Investment (FDI), Economic growth (GDP), Panel co integration, Panel causality*

1   Introduction

The relationship between Foreign Direct Investment (FDI) and economic growth is extensively examined subject in the developing economies, both empirically and theoretically; and still remains a much debated topic. In the past two decades the pattern of FDIs shows a continuous increasing trend. FDIs still remain an important tool for developing economies to enhance growth. According to Shaari et al (2012) countries that have weaker economy consider FDIs as the main and only source of growth as well as economy modernization. Thus, governments of developing economies are constantly focusing on foreign investments.

The reason that this investigation focuses in the South Eastern European Economies is due to the similar market characteristics (i.e. growth, unemployment, exports etc.).

The fact that countries covered in the sample come from communist regime, and after the breakdown of the regime we have witnessed a mass privatization process; mainly coming through FDIs. Nevertheless, these countries are intensively adopting aggressive strategies and policies to attract FDI. Yet, the aim of the investigation is to

provide an insight of the impact of FDI in economic growth. More precisely the aim of the research is to identify the causal effect among the variables: whether FDI causes economic growth or vice versa. The contribution of the research will be significant and helpful for policy makers in the respective countries. The outcome will provide facts for policy makers of whether economic growth causes FDI or it is FDI that causes growth. Nevertheless, bidirectional causality may occur as well.

The role FDI favors the transition economies of Central and Eastern European Countries, yet some critical views have highlighted that capital inflow effects have been uneven and uncertain (Pavlinek and Smith, 1998 and Szanyi, 2000. In addition, (Sharp and Barz 1997, and Castellani and Zanfei 2006) point out that FDI can create potential for enhancing industrial production, productivity, modernization, enhancement in quality and competitiveness output. This outcome has been highlighted as a result of organizational restructuring, workers training, technological transfer, inflow of management practices, and modernization of production and organization strategies. Moreover, Blostrom and Kokko (2003) state that FDI not only has the potential to increase host a country's' economic growth and reduce unemployment, yet it goes far beyond this by enabling the host country to have access to more advanced technology and "know how". FDI is closely interrelated with production efficiency (De Mello, 1999). This is due to foreign technological spillovers in the host country and the increase knowledge of the workers. With regards to the benefits that the host country gains; Southeastern European economies have shown tremendous interest to enhance FDI inflow in the past two decades.

Contrary to the above researchers, Heinsz (2003) emphasizes that FDI does lead to industrial improvements, modernization, increase in productivity, enhancement in quality and competitiveness, integration in capital markets, yet all these characteristics represent benefits to FDIs, excluding economic growth of the particular host country. Moreover, Buckley et al (2002) argues that the extent to which FDI will contribute to the host country highly depends on the social conditions of that country.

2   Literature Review

The relationship between FDI and economic growth has been significant in many transition economies. Yet each study suggests different results, depending on the area of the study. Due to different conclusions FDI and economic growth remains a big debate over the impact that the former has on the economy. Ozturk and Kolyonzu (2007), perform empirical investigation based on cross-country comparison.

The countries chosen are Turkey and Pakistan over the period 1975-2004. In order to analyze the causal relationship between FDI and economic growth they perform the Engle-Granger cointegration and Granger causality tests. They find that those two variables are cointegrated for both model countries (Turkey and Pakistan). Their empirical findings suggest that it is GDP that granger causes FDI for the case of Pakistan while on the other hand; Turkey shows a strong evidence of bidirectional causality between the two variables.

Adraz and Rodriguez (2010) investigate the causal relationship between exports, inward FDI and economic growth for the case of Portugal. They consider the period from 1977 to 2004. Their conclusions suggest that FDI does foster economic growth in the long- run while in the short run they suggest that there is a bi-directional relationship between the variables. Moreover, Tang and Selvanathan (2008) examined the causal link between FDI, domestic investment and economic growth for the case of China between the periods 1988 – 2003. Their technique is based in the multivariate VAR and ECM. Their results suggest that there is bi-directional causality between domestic investment and economic growth; however, there is single direction causality from FDI to economic growth. Furthemore, a study by Tao et al (2009), investigate the relationship of FDI and economic growth for the case of Shanghai. Their investigation is based on the time series data from 1981-2008. They use Augmented Dickey Fuller (ADF), followed than by granger causality test, and multivariate regression model. The results show that FDI granger causes GDP for the case of Shanghai.

Liu et al (2002) examined the long run relationship between economic growth foreign direct investment and trade for the case of China. He used the cointegration framework in quarterly basis. His investigation found a bi-directional relationship among FDI, economic growth and exports for the period 1981 – 1997. Further studies for the case of China are done by Wang (2002). He explores the kinds of FDI inflow that contribute on economic growth considering the period from 1987 – 1997. She suggests that manufacturing FDI's are the ones that cause economic growth only.

While many authors find out that FDI has a positive relationship with economic growth of a particular country and the relationship is FDI causing economic growth there is evidence that proves the opposite. Borensztein et al (1998) examined the empirical relationship between FDI and economic growth for the developing economies. They suggest that FDI has allowed these countries to transfer technology and higher growth in cases when the host countries had a minimum threshold stock of human capital. Moreover they conclude that the main reason that FDI promotes economic growth is by increasing technological progress in the host country, unlike increasing total capital accumulation in the host country. In their investigation they used data of gross FDI which is referred only to inflows, while for the economic growth they used the growth rate on income as the annual average rate of GDP per capita. Their conclusions suggest that in countries with low level of human capital the effect of FDI on economic growth is negative, while the causality runs from GDP to FDI.Willawera et al (2010) examine the relationship by using a panel of data for 45 countries from year 1997 to 2004 and proves the opposite. The results generated show that FDI inflows do contribute to economic growth GDP in the presence of highly skilled labor, however; in countries with higher rates of corruption the relationship between FDI and economic growth GDP is negative.

Further studies by Duttaray et al (2008), show that FDI causes growth in several developing countries in a sample of 66 developing countries. However, from country to country the mechanism differs and reverses causality from growth to FDI holds for 20 countries, and 44% of the countries, FDI enhance economic growth.

Orman and Bolbol (2003) explain some of the reasons. They conducted a pooled OLS for 17 Arab countries they find that the impact of FDI on economic growth

largely depends on the development phase of the financial sector. While, Bailliu (2000) suggests that capital inflows are capable to having an impact in economic growth only in countries with sufficiently developed financial sectors.

Even though there is extensive research in this topic, still, the causal relationship among the variables remain very much debated. Due to this fact, there is plenty of room for further investigations. The methodology that will be used in this research will try to give an answer of the direction of the causality, of whether FDI causes economic growth, economic growth causes FDI, or there is bi-directional causality.

3   Methodology

The methodology that will be applied in conducting this particular study; with regards to FDI and economic growth for the case of Southeastern European economies (Croatia, Serbia, Bosnia and Herzegovina, Montenegro, Albania, Kosovo, and FYROM-Macedonia) will be performed in three steps. First step, test for examining the order of integration will be employed for the FDI and GDP time series. Since the time period of examination is very short, will be employed the recent developed panel unit root technique in order to increase the validity of the study. Second step, after establishing the order of integration of the series panel cointegration test will be performed in order to examine the long run relationship of the variables. Third step, dynamic heterogeneous panel causality will be performed to assess the short run cointegration. Finally, the direction of causality between variables will be examined by using heterogeneous panel causality test.

3.1 Order of Integration

Order of integration is a methodology used that reports the minimum number of differences required to obtain a covariance stationary data. Since most of the macroeconomic data are non-stationary (they have unit root), tests for stationary have to be performed. Further, the cointegration technique applies to series that are not stationary and it requires to that variables (FDI and GDP) to be integrated of the same order I(1). Two common tests can be performed for this purpose: Augmented Dicket Fuller (ADF) as well as, Philips Perron (1988). Tests will be performed with considering the trend and intercept.

3.2 Engle and Granger (1987)

Engle and Granger (1987) tests will be performed after it is checked for integration and cointegration properties of the series. The Engle and Granger technique is a useful tool to test for causality of the variables (i.e. GDP causes FDI or vice versa). The standard Granger causality test (1969) do not take into consideration the existence of the long-run relationship of the variables, yet: this test is efficient in capturing the short-run relationship.

4   Conclusion

Southeastern Economies are constantly formulating strategies to attract FDIs. Yet, no studies are done with regards to the impact that FDI has on economic growth. Moreover, the fact that there is vague condition in the causal relationship among FDI and growth, there is no clear cut answer that policymakers can provide with certainty.

This study aims to explore the impact of FDI and economic growth in the selected Southeastern Economies (Croatia, Serbia, Bosnia and Herzegovina, Montenegro, Albania, Kosovo, and FYROM-Macedonia) for the period 2000-2013. Vast majority of the studies suggest that there is a relationship between FDI and economic growth; running from FDI. Nonetheless, there are cases of economic growth causing FDI as well as, bi-directional relationship between the variables.

The study aims to investigate the causal relationship between FDI and economic growth. The research is based on proven methodologies by applying Engle and Granger (1987) technique in order to capture the short-run and long-run dynamics of the variables.

References

Andraz, J. & Rodrigues, P. (2010) What causes economic growth in Portugal: exports or inward FDI?.*Journal of Economic Studies*. p. 267-272.

Bailliu, J.N. (2000) Private capital flows, financial development and economic growth in DevelopingCountries. *Bank of Canada Working Paper no. 2000-15*, Bank of Canada, Ontario. P.12-21

Blostrom, M. and Kokko, A. (2001) Foreign Direct Investment and Spillovers of Technology. *International Journal of Technology and Management*, p. 435-454.

Buckley, P. Clegg, J and Wang, C. (2002) The impact of inward FDI on the performance of Chinese manufacturing firms," *Journal of International Business Studie.* p. 637–655.

Borensztein E. De Gregorio J. and Lee J.W. (1998) How does Foreign Direct Investment Affect Economic Growth? *Journal of International Economics.* p.115-35.

Carkovic, M., and Levine, R. (2003) Does foreign Direct Investment Accelerate Economic Growth*? University of Minnesota Department of Finance Working Paper*

Casson, M. (2007) Multinational Enterprises: Their Private and Social Benefits and Costs. *World Economy*, *30*(2). p.308-328.

Castellani, D. and Zanfei, A. (2006) Multinational Firms, Innovation and Productivity. *Cheltenham: Edward Elgar*. p.73-98

De Mello, L. R. (1999) Foreign Direct Investment: Empirical Evidence on Eeffects and Alternative Policy Options. *Canadian Tax Journal*. p.1139-1180.

Duttaray, M.D. and Mukhopadhyay, K. (2008) Foreign Direct Investment and Economic Growth in Less Developed Countries: an Empirical Study of Causality and Mechanisms. *Applied Economics*. P. *40-50*

Hansen, H. & Rand, J. (2004) On the Causal Link Between FDI and Growth in Developing Countries. *Institute of Economics University of Copenhagen. P. 30-35*

Heinsz, W. J. (2003) Concentrated Power, foreign Direct Investment and Economic Growth. *Wharton School of Business*. p.25-42

Liu. V. and Suruga, T. (2005) Foreign direct investment, public expenditure and economic growth: the empirical evidence for the period 1970-2001, *Applied Economics Letters*. p.45 – 49

Engle, R.F. and Granger, C.J (1987) Co-integration and Error Correction Representation, Estimation and Testing. *Econometrica* 55(2) p.251-276

Mossa T. H. (2002) Foreign Investment and Development: The New Policy Agenda for Developing Countries and Economies in Transition. *Washington, DC: Institute for International Economics journal* . p.12-31

Omran, M. & Bolbol, A. (2003) Foreign Direct Investment, Financial Development and Economic Growth: Evidence from the Arab Countries. *Review of Middle East Economics and Finance*. p231-249.

Ozturk, R. and Kalyoncu, H. (2007) Foreign Direct Investment and Economic Growth: An Empirical Investigation Based on Cross-Country Comparison. Cag University. p.73-78

Pavlinek, P. and Smith, A. (1998) Internationalization and Embdedness in East Central European Transition: The Contrasting Geographies of Inward Investment in the Czech and Sllovak Republic. *Regional Studies.* p.619-638

Philips, P.B. and Perron, P. (1988) Testing for a Unit Root in Times Series Regression. Biometrica. p.335-446

Rehman et al (2010) Impact of Foreign Direct Investment (FDI) inflows on Equilibrium Real Exchange Rate of Pakistani. *Reasxearch Journal of South Asian Studies*. p.35-41

Shaari, B. H., Hong, H.T. and Shukeri, N.S. (2012) Foreign Direct Investment and Economic Growth: Evidence from Malaysia. *International Business Research*. p.100-106

Sharp, M.and Barz, M. (1997) Multinational Companies and the Transfer of New Technological Capabilities in Central and Eastern Europe and the Former Soviet Union. In: D.A. Dyker, ed. The technology of transition: Science and technology policies for transition countries. Budapest: Central European. *University Press*, p.95–125

Tao I. et al (2009) Empirical Research on the Effect of FDI on Economic Growth-Evidence from Shanghai. *Management and Service Science*. p.37-68

Thompson, A. (2008) FDI and Economic Growth in Malaysia. *Journal of Applied Econometrics*. p.78-83

Wijeweera, A. Villano, R. & Dollery, B. (2010) Economic Growth And FDI Inflows: A Stochastic Frontier Analysis. *Journal of Developing Areas*. p.143-158.

Role of Fertilizer subsidy policy in Agricultural Development. The case of Sri Lanka

Kodikarage N. Nadeeshani Silva[1] and Dieter Kirschke[2]

[1]Department of Agricultural Economics, Faculty of Agriculture, University of Ruhuna, Sri Lanka
[2]Division of Agricultural Policy, Faculty of Agriculture and Horticulture, Humboldt University in Berlin, Germany
nadeeds@gmail.com

Abstract. This research article examines the impact of fertilizer subsidy programme on paddy cultivation in Sri Lanka since 1980 to the date. Fertilizer subsidy programme has been playing a dominant role in Agriculture development in Sri Lanka since 1962 as most long lasting politically sensitive policy programme. At present fertilizer, subsidy policy is criticizing due to its Economic, social, political and Institutional implications in addition to increasing burden to the government. The study has attempt to analysis the impact of fertilizer subsidies with reference to Paddy cultivation in Sri Lanka. This issue has analyzed employing Demand functions for major fertilizer types. Research outcome has further elaborated using descriptive statistical methods supporting to Secondary data. Results have shown that, changes in the prices of fertilizer and paddy do not have a significant effect on fertilizer usage, which points to the fact that the fertilizer subsidy is not a key determinant of the use of fertilizer in paddy cultivation. Further, paddy yield has positive relationship with Urea fertilizer use. Self- sufficiency in rice is being increasing in country with the incentive of growing more paddy land rather than use of fertilizer. Rural household welfare has affected by reducing cost of production and thus enhancing farmer's income. Further, study has concluded that fertilizer subsidy could withdrawn with time, replacing effective farmer supportive programme ,with the incentive of organic fertilizer usage and developing appropriate infrastructure and institutional facilities that require to increase productivity in paddy cultivation.

Keywords: *Fertilizer subsidy, Paddy Farming, Economic Impact, Demand Function, Social welfare*

1. Introduction

In this research, fertilizer subsidy policy in Sri Lanka has evaluated using secondary data of paddy sector which is main beneficial crop of Fertilizer subsidy programme. Rice is main staple as well as main economic crop in Sri Lanka. Rural economy in Sri

Lanka heavily depends on paddy sector. Government of Sri Lanka has introduced many policies and programmes to increase paddy production since independence in year 1948. The fertilizer subsidy program is one of the longest lasting, most expensive and most politically sensitive policy implemented to promote rice cultivation in Sri Lanka (Weerahewa et al, 2010). Fertilizer subsidy programme in Sri Lanka has been continuing in Sri Lanka since 1962 with the onset of Green Revolution, later with some structural changes.

Background Information

The main objective of fertilizer subsidy policies was to encourage paddy farmers to cultivate high yielding rice varieties, which are highly responsive to chemical fertilizers. Regardless of the significance impact of subsidy policy on rural Agriculture sector, it has become a huge financial burden to Sri Lankan government. Anyhow, fertilizer subsidy has become customary in every successive government, continuing subsidy schemes except short pause period from 1990-1994. According to Central Bank of Sri Lanka (2012), Fertilizer subsidy payment constitutes 2.24 percent of total government expenditure. Cost of import has increased with the cost of inorganic fertilizer import mainly from China and United Arab Emirates. Even though Sri Lanka has its own resources, it has unable to extract as in good quality due to technological limitations.

Many studies have revealed failure and weakness of fertilizer subsidy programme regardless of the positive impacts of the Fertilizer subsidy programme. It has widely accepted that land productivity improvement led by fertilizer subsidy programme due to induced motivation of farmers to expand the land under paddy cultivation (Weerahewa et al 2010). Other than the projected social and economic benefits of Fertilizer subsidy programme, Sri Lankan socio economic and environmental condition has been negatively affected. The fertilizer subsidy policy and its implications have directly been associated with the Fertilizer use and indirectly with Production level, changes of rice and fertilizer price level, changes of food consumption and changes of social welfare.

This study has addressed the role of Fertilizer subsidy policy in Agriculture Development and subsequent impact, challenges associated with Policy intervention in context to Sri Lanka. Specially, the fertilizer subsidy policy has been created Economic, social, political and Institutional implication from the beginning of Fertilizer schemes. Expected output of the Fertilizer subsidy has deviated from its basic target with the limitation of other resource limitations; land, technology and other input constraint (Ekanayake, 2006). Further, Political power in Sri Lanka mainly depends on the votes of rural population, which represent most influential group in Sri Lankan general election. Mainly, Paddy cultivation provides livelihood opportunities for more than 1.8 million farmers in country and any significant deviation from the status quo could damage the political power base of the present country ruling government (Weerahewa et al, 2010). Prevailing administrative and bureaucratic agenda of the Fertilizer subsidy schemes has to improve in order to develop agriculture sector in Sri Lanka. Based on literature review and background information of the research problem, main research question of the study was; what are the major policy implications of the Agricultural

Fertilizer subsidy policy in Sri Lanka in context to the Economic, Social and Institutional aspects?

2. Literature review

Literature finding has collected to underpin the empirical evidence through theoretical background of the input subsidy policy. Fertilizer subsidy policy considered as variable input policy. Major instruments of input policies used by government have three main dimensions; Price Influence, Interventions in the delivery of farm inputs, Provision of Information on inputs to farmers (Ellis, 1992). There are many instruments used for provision of fertilizer subsidy as a fixed subsidy or as a variable subsidy. Following figure, 01 shows the equilibrium in the fertilizer market with a fixed and a variable subsidy at a theoretical level in more detail.

Suppose that Pw is the world market price of fertilizer and a fixed subsidy given to fertilizer importers at a level of s, in which case the retail prices in the domestic market will vary depending on the variations in the world market prices. In such a context, the price of fertilizer in the retail market will be (Pw – s) and the quantity imported at this price will be F. If the government continues to provide the subsidy at a level of s, regardless of the world market price of fertilizer, when the world market price falls to Pn the retail price of fertilizer would be (Pn – s). At this price, Fn units will be imported. Alternatively, the government could fix the retail price at Pw – s, in which case the level of subsidy will have to adjust depending on the world market price. In this situation, the quantity of fertilizer demanded and imported will be equal to F regardless of the price of fertilizer in the world market. The level of the variable subsidy at a world market price of Pw will be s, and the level of the variable subsidy at a world market price of Pn will be s´. The variable subsidy approach does not allow the local fertilizer market to adapt to changes in the world market and provides a predictable environment for farmers.

When the world market price is Pw, the cost of the subsidy to the government would be F · s regardless of whether the subsidy is variable or fixed. When the world market price falls to Pn with a fixed subsidy of s, the cost of the subsidy would be Fn.s
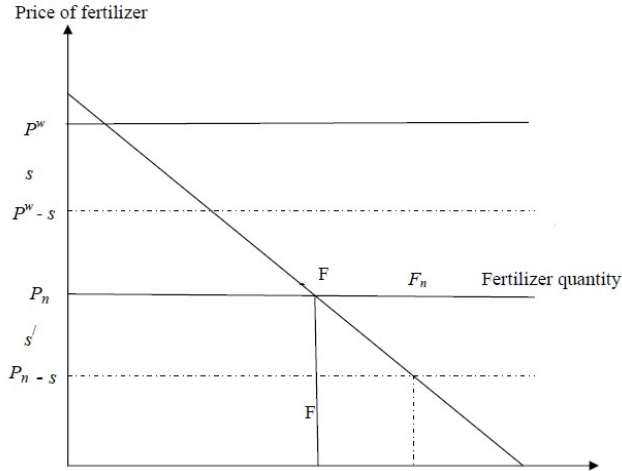
Fig.1. Equilibrium in Fertilizer market and subsidy issue (Source: Weerahewa , 2004)

Elasticity of Demand and supply

With elastic demand for a certain input, providing a small subsidy on the input may significantly increase its use. Price elasticity of supply (demand) measures how much the supply (demand) of a good (or service) in the market changes when the price of the good changes by one percent. For a market good with a supply elasticity of one, a one percent rise in price results in a one percent increase in the quantity supplied. Similarly, for a good with a demand elasticity of one, consumers will demand one percent less of the good when its price increases by one percent. Elastic (Inelastic) supply or demand indicates that quantity of the good supplied or demanded changes greatly (little) when the price of the good changes slightly.

Elasticity can calculated for changing price ratios, fertilizer/ price ratio. Accordingly, demand of fertilizer has elastic nature. It is important to notice that, the elasticity of demand for inputs generally will depend on both the prevailing production technology used and crop prices.
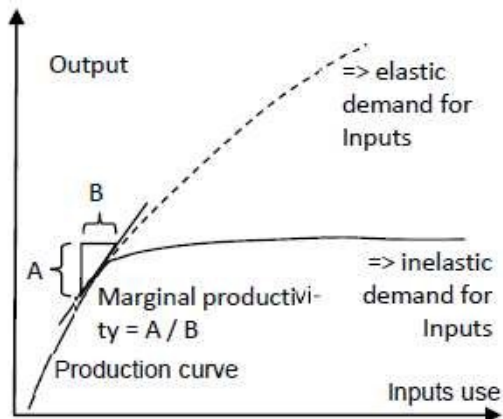
Fig. 2. Marginal Productivity and Elasticity of Demand (Source: Source: Andrew and Chirwa, 2009)

Inefficiency of fertilizer subsidy programme

Economic theory has explained that if farmers are unable to use those subsidized inputs, it will create a deadweight loss to society. A deadweight loss can defined as a loss of economic efficiency resulting from the misallocations of resources where those who extract less value from the goods than their cost, when their cost is valued at non-subsidized prices (Takeshima and Lee, 2012). The following figure 03 illustrates a simplified example of the potential deadweight loss that arises from a fertilizer subsidy.
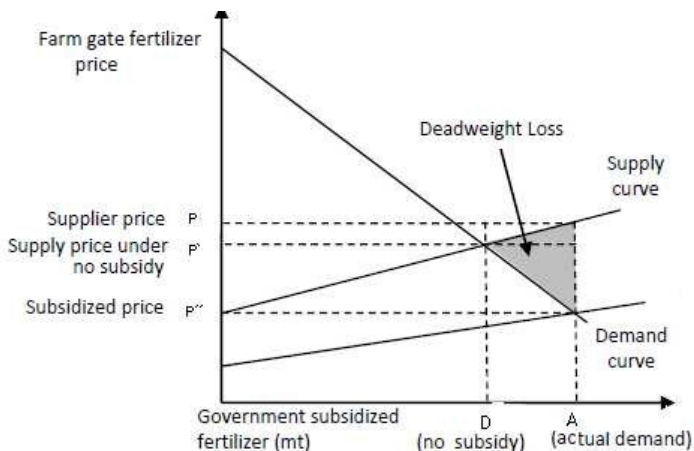
Fig.3. Deadweight loss of Fertilizer subsidy programme (Source: Hiroyumi and Lee, 2012)

The reason for inefficiency of input subsidies may arise due to several reasons. Public resources that use for supplying fertilizer subsidy have associated opportunity cost.

these grants can as well be used in many other ways which might provide greater social benefits, such as research and development, reducing tax burden on individuals and firms to promote increased investment into productive assets, direct income transfers to the poor in order to help them accumulate productive assets, increase productivity, and increase their income (Hiroyumi and Lee, 2012).

Impact of fertilizer subsidy programme

As emphasized in many research studies, Fertilizer subsidy scheme have economic, social and ecological impact in a way that giving both positive and negative impacts. Nevertheless, the major and direct impact of fertilizer subsidy programme is to increase the crop yield and thus increase productivity. Narul, et al (2012), Weerahewa et al (2010), Andrew and Chirwa, (2011), IDS studies (2008) has empirically estimated those direct production impact of fertilizer subsidy in Sri Lanka.

Further, fertilizer subsidy has secondary effect: a).Production impact by increasing production, increasing self-sufficiency, b). Market impact changing price and import and export market, c). Impact on Macro Economic Impact (Real GDP impact, inflation etc), d). Impact on social welfare and e). Environmental impact. Removal of fertilizer subsidy has a negative impact to the rice farming in Malaysia by decreasing rice Production and self-sufficiency level and increasing rice imports (Narul et al, 2012).

Input Market impact due to subsidy programme is critical to different suppliers and this has studied by Andrew and Chirwa (2011) shown that incremental sale of fertilizer due to subsidy programme. It has quantified that in year 2005, which was starting year of programme the sales was 60 % while it has increased up to 90 % in year 2009. Fertilizer use has increased tremendously and nearly double of the production in year 2009 compared to 2005.

3. Data and Methodology

This study has based on organizing, collecting and analyzing secondary data scattered in several government publications, statistical reports, farm and household reports and agriculture sector reviews. Publications of Central Bank of Sri Lanka, Department of Census and Statistics and Department of Agriculture were major data sources. Main objective of the study was to analysis the implications of fertilizer subsidy policy in Sri Lanka in context to Economic, Social and Environment and Institutional aspects. Following conceptual framework was used to analysis impact of fertilizer subsidy programme.

Table 1. Conceptual framework

Impact	Variable	Expected outcome
1. Direct impact	Fertilizer demand	Increase

2. **Indirect impact**	Production (yield, land extend, self sufficiency in rice)	Increase
	Market impact (Paddy farm gate price, fertilizer price, rice consumer price)	Increase/Decrease
	Social impact (Expenditure on Subsidy, cost of production)	Increase/Decrease

In order to determine the fertilizer demand for main three type of fertilizer, following analytical tool and methods used. Factors affecting to fertilizer demand assumes as,

1. Fertilizer Prices (Rs. /Mt)
2. Paddy Prices (Rs. /Mt)
3. Extent under irrigation schemes (Ha)
4. Policy changes

Three main fertilizer functions used based on data available from year 1980-2009 covering all four phases of fertilizer subsidy programme. Regression models specified for annual use of fertilizer in paddy farming at the national level used as the dependent variable and current fertilizer nominal farm gate price and paddy land extent as independent variables. Major assumptions for the model and analysis were that all other factors that influence fertilizer use were remained unchanged. Dummy variables used to interpret policy changes. Further, it also assumed that total amount of different fertilizers issued for paddy cultivation during the given year utilized fully for paddy cultivation.

Own price and cross price elasticity of demand has calculated from estimated demand function and those elasticity explain the sensitivity (elasticity) of fertilizer demand to change in fertilizer prices and paddy prices, which implies the impact of fertilizer subsidy policy on paddy cultivation.

Description of Econometric tools used

Augmented Dickey- Fuller (ADF)

Before estimating models, it is essential to verify whether the series analyzed are stationary or not and whether there are one or several long run equilibrium relationships between them. If a series is stationary means, its variance and auto covariance are independent of time. In order to test stationary and co-integration among the variable in the model multiple regression and the Augmented Dickey- Fuller (ADF) test has done.

Regression Test

The natural logarithm values of all the variables used to estimate fertilizer demand function by using simple linear regression model. The model is as follows:

$$Qt = f \{(P \ (f) \ (t)), (P \ (p) \ (t\text{-}1)), (E \ (p) \ (t)), (D)\}$$

Qt – Quantity of fertilizer (Mt) used in year t

P (f) (t) – Price of Fertilizer (Rs. /Mt) in year t

P (p) (t-1) – Farm Gate price paddy (Rs. /Mt) in year t-1

E (p) (t) – Paddy Extent under Irrigation in year t

DU- Policy changes

4. Empirical analysis

Economic implication of fertilizer subsidy programme

Impact on fertilizer consumption and paddy yield

Most direct and immediate impact of the any input subsidy policy is to increase the fertilizer consumption. This impact has been evaluated using both descriptive and econometrics analysis. Changes of fertilizer consumption during 1980 to 2009 have shown in following Fig 04.
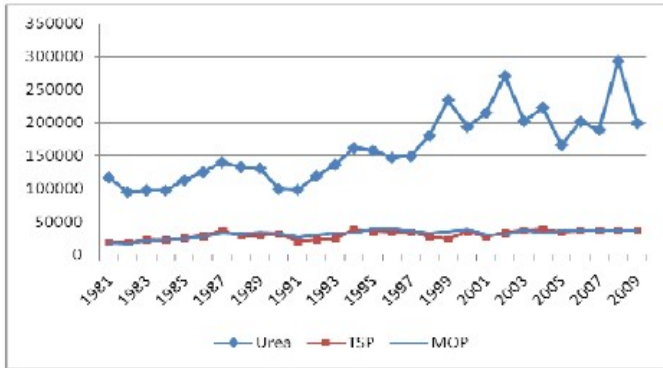
Fig.4. Fertilizer consumption in Paddy farming in Sri Lanka (Source: Based on calculation made using data from central B ank Annual reports and Publications of National fertilizer secretariat, Department of census and statistics and Department of Agriculture in Sri Lanka, Various issues.)

Above Fig. 4 has shown that there is a noticeable increment in Urea usage in paddy farming while use of TSP and MOP remained the same level over time. Nevertheless, fertilizer use in paddy sector has increased from 178,609 Mt in 2006 to 415,403 Mt in 2012 (Central Bank, 2012).

In years 1991,1992, and in 1993, the overall consumption of fertilizer increased by 3.5 percent, 16.0 percent and 17.0 percent respectively indicating that most of the farmers were gradually adjusting themselves to the new prices. Use of fertilizer for paddy

recovered in 1998 due to the increase in usage of highly subsidized Urea by 20.6 percent over that of 1997 (Ekanayake, 2006). In 1999, use of Urea for paddy increased substantially to a record level, indicating that paddy farmers benefited most from the fertilizer subsidy scheme. In year 2000, the farm gate price of paddy declined and thus farmers tend to cultivate less land extend leading less fertilizer consumption. Variations of total fertilizer consumption varied with adverse climate condition, farm gate price of paddy and Fertilizer prices.

Fertilizer Demand Function

Demand function for urea

Dependent variable: Urea use
Independent variables: Urea Price, Paddy farm gate price, Land Extend, Policy change

Table 2. Multiple Regression analysis on Urea use

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	-36774.3	59762.2	-0.615344	0.5441
UP	-1.14253	0.430526	-2.65379	0.0139
PP	10.4209	1.70451	6.11369	0.0000
EX	146.474	73.4996	1.99285	0.0578
DU 1	-34636.5	14567.4	-2.37767	0.0257

R-squared = 73.9751 percent, R-squared (adjusted for d.f.) = 69.6376 percent
Standard Error of Est. = 29239.1, Mean absolute error = 19917.7,
Durbin-Watson statistic = 1.47718 (P=0.0225)
Urea use = -36774.3 - 1.14253*UP + 10.4209*PP + 146.474*EX - 34636.5*DU 1

The entire variable were found to be stationary all first difference and it was concluded that all the series were I (I). According to above table, there is a statistically significant relationship between the variables at the 95.0% confidence level. Based on the Durbin-Watson (DW) statistic there is an indication of possible serial correlation.

Further, elasticity of each independent variable to the Urea consumption could expressed through estimated values. All independent variable except Paddy land extends are significant for urea usage with theoretically expected signs. Further, these results indicate that fertilizer demand is inelastic to own price and Policy changes. Nevertheless, it is elastic to paddy price and Land extend. However, this model seems to be inappropriate to elaborate the policy implication on Urea consumption. However, Ekanayake (2006) has shown that, fertilizer demand is inelastic to own price, cross price and policy changes.

Demand Function for TSP use

Dependent variable: TSP use
Independent variables: TSP Price, Paddy farm gate price, Land Extend, Policy changes

Table 3. Multiple regression analysis on TSP use

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	30768.5	14747.6	2.08634	0.0500
TSP P	0.264924	0.501315	0.528457	0.6030
PP	0.470069	1.13317	0.414827	0.6827
EX	-9.68572	16.5549	-0.585068	0.5650
DU 2	-1728.3	2887.22	-0.598603	0.5562

R-squared = 33.9694 percent, R-squared (adjusted for d.f.) = 20.7633 percent
Standard Error of Est. = 5731.5, Mean absolute error = 4333.46
Durbin-Watson statistic = 1.34816 (P=0.0148)
TSP use = 30768.5 + 0.264924*TSP P + 0.470069*PP - 9.68572*EX - 1728.3*DU 2
 Based on the Unit root test, the entire variables were found to be stationary at first difference and it was concluded that all the series were I (I). According to above tables, there is not a statistically significant relationship between the variables at the 95.0% confidence level. Based on the Durbin-Watson (DW) statistic there is an indication of possible serial correlation at the 95.0% confidence level.
 Based on the statistical analysis, none of the independent factors is significant for TSP usage with theoretically expected signs. Nevertheless, this model does not show significant relationship with TSP Consumption with other independent variable thus do not show significant value to measure elasticities.

Demand Function for MOP use

Dependent variable: MOP use
Independent variables: MOP Price, Paddy farm gate price, Land Extend, Policy changes.

Table 4. Multiple regression analysis on MOP use

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	15659.8	13659.7	1.14642	0.2652
MOP P	-0.600815	0.484215	-1.2408	0.2290
PP	2.51234	1.1079	2.26765	0.0346
EX	0.331218	14.9689	0.0221271	0.9826
DU 3	2636.28	2328.23	1.13231	0.2709

R-squared = 48.3965 percent, R-squared (adjusted for d.f.) = 38.0758 percent
Standard Error of Est. = 5081.56, Mean absolute error = 3973.26
Durbin-Watson statistic = 1.33162 (P=0.0141)

MOP use = 15659.8 - 0.600815*MOP P + 2.51234*PP + 0.331218*EX + 2636.28*DU

Based on the Unit root test, the entire variable were found to be stationary all first difference and it was concluded that all the series were I (I). According to above table, there is not a statistically significant relationship between the variables at the 95.0% confidence level. Based on Durbin-Watson (DW), there is an indication of possible serial correlation. Result shows that, except paddy price, other factors are not significant for MOP usage with theoretically expected signs. Further, these results indicate that fertilizer demand is inelastic to own price and paddy land extend. However, it is elastic to paddy price.

Production impact

Yield response of paddy for Urea
   Concerning yield response of Urea usage of paddy in Sri Lanka during last few decades, it has clear relationship between yield responses with the use of Urea use. Data from National Fertilizer Secretariat show that the paddy sector used about 53 percent of the fertilizer supplied in 2006 (Weerahewa et al. 2010). Paddy sector has used 132,800 metric tons of the 252,800 metric tons of total fertilizer supplied in 2006 for whole crops.
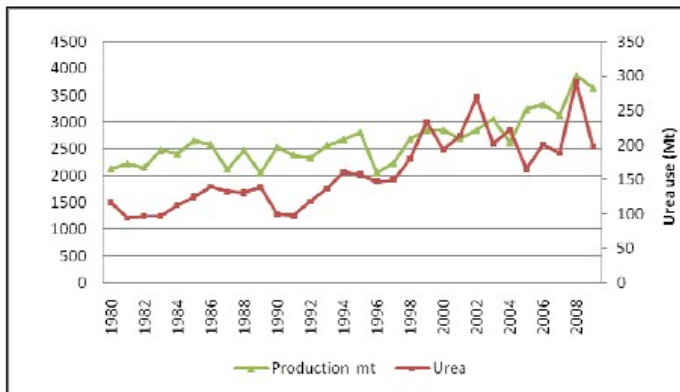
Fig.5. Paddy Yield response for Urea fertilizer (Source: Based on calculation made using data from Publications of National fertilizer secretariat, Department of census and statistics and Department of Agriculture in Sri Lanka, Various issues)

   Fertilizer application rates more than doubled among rice farmers over the period 1983-2005 and average yield increased by only about 9 percent (Tibbotuwawa, 2010). The variability of agro ecological zones has an impact on paddy yield in spite of the subsidy programme (Chandrasiri and Karunagoda, 2008). Regional differences in

technical efficiency of fertilizer use and 32 percent increase in fertilizer use (because of changes in the subsidy scheme) resulted in a 17 percent yield increase from 2005 to 2008.

Cultivated land extend

Following figure, 06 shows the paddy land extend under paddy cultivation over last two decades. It shows that significant increase in land extend under paddy cultivation after very recent fertilizer subsidy programme which was target mainly paddy farmers.
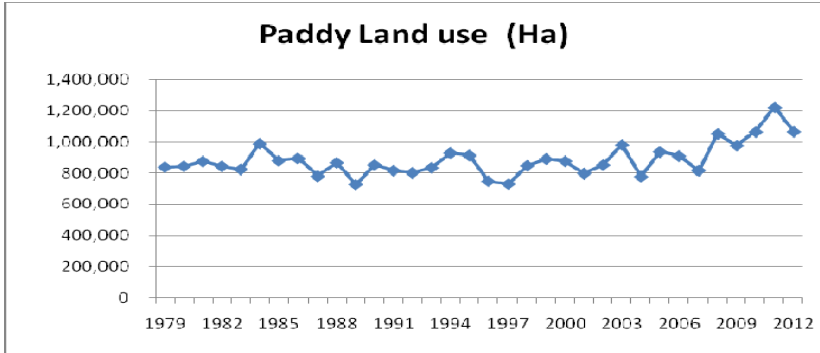
Fig. 6. Land Extend under Paddy cultivation (Source: Based on calculation made using data from Publications of National fertilizer secretariat, Department of census and statistics and Department of Agriculture in Sri Lanka, Various issues)

Self-sufficiency of Rice

One of the major objectives of the fertilizer subsidy programme in Sri Lanka is to sustain the country by achieving self-sufficiency in Rice. Therefore, it has considered as one of the major production impact. Following fig. 7 shows the self-sufficiency in rice in Sri Lanka since 1970 to 2010.
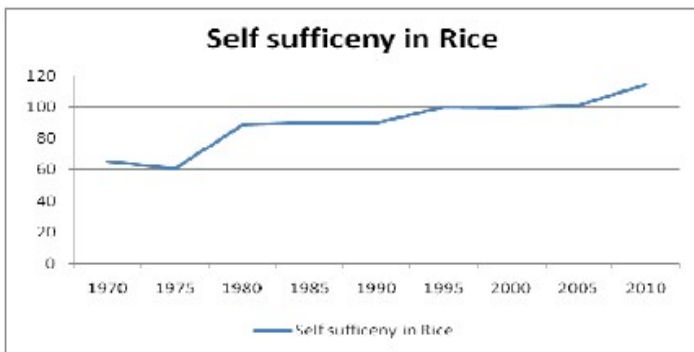
Fig.7. Self-sufficiency in Rice in Sri Lanka (Source: Self calculation using data from Department of Agriculture; Department of Census and Statistics, various years)

Concerning self-sufficiency in rice in the country, there is a noticeable increment of self-sufficiency within country and thus rice import has decreased over the time. This is a one positive achievement of the country in context to objective of the fertilizer subsidy programme.

Market Impact

Market impact due to fertilizer subsidy has analyzed through changes of retail price of rice, paddy farm gate (Producer price) and Fertilizer retail market price since those are most important variable to analysis market impact of fertilizer subsidy programme.
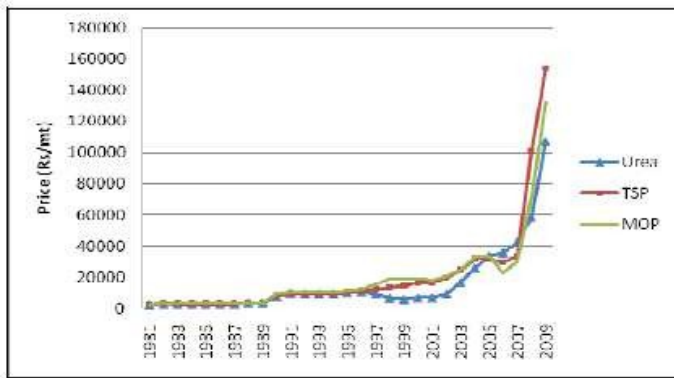
Fig. 8. Market Price changes of subsidized Fertilizer (Source: Based on calculation made using data from Publications of National fertilizer secretariat, Department of census and statistics and Department of Agriculture in Sri Lanka, Various issues)

In early 1980, all fertilizer prices reduced sharply since the government increased its subsidy rate from Uniform subsidy of 50 percent of cost and Freight (C & F) fertilizers to 85 percent and 75 percent of the C&F price of Urea, MOP, and TSP respectively. During the period of 1991 to October 1994, fertilizer prices remained stable after completely withdrawn the fertilizer subsidy in year 1990. The sale price of Urea maintained at the indicative price under the fertilizer subsidy scheme. Prices of other fertilizers were determined by the market considering the factors such as appreciation of US$, price fluctuations in the international market and freight rates. (Ekanayake, 2006).

Retail price of major rice type in Sri Lanka
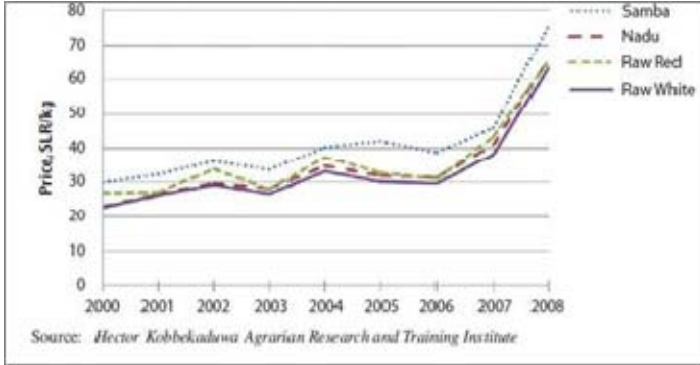
Fig. 9. Retail Price changes of major rice varieties (Source: Self calculation using data from Department of Agriculture; Department of Census and Statistics, various years)

According to above Fig. 9 there is a general increment of selected rice varieties in Sri Lanka. Further, it has shown very clearly, the unique pattern of changes of all rice varieties. Noticeably, after most recent subsidy programme, initially there has been a price reduction and later continues price increment may be due to other market factors.

Paddy farm gate price
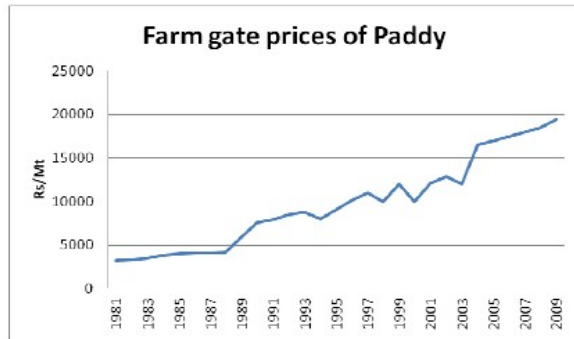
Fig. 10. Changes of farm gate price of Paddy (Source: self calculation using data from Department of Agriculture; Department of Census and Statistics, various years)

Paddy price has heavily fluctuated during the study period. Among the most remarkable years of study, in 1998, there were high paddy productions, which lead to decrease average producer price over the previous year. However, this price remained well above the guaranteed price of paddy. In contrast to the normal behavior, in 2000 the farm gate price of paddy declined mainly due to rice importation and drop of paddy production. In 2001, farm gate price increased and maintained at a higher level until 2003. Paddy price dropped in 2003 with the higher production.

Social impact of fertilizer subsidy

Social impact may include the social government cost for fertilizer subsidy, which has an opportunity cost for use for other social welfare activities including infrastructure development, educational and health improvement programme. Following table 05 shows the Public expenditure on fertilizer subsidy and as its share of total government expenditure.

Table 5. Fertilizer Subsidy as a Percentage of total Government expenditure (Source: Self calculation using data from Central Bank of Sri Lanka, Department of Census and Statistics, various years)

Year	Expenditure on Fertilizer subsidy (Rs. Millions)	Expenditures on fertilizer subsidy as a % of total government expenditure
1998	2,152	0.8
1999	1,390	0.5
2000	1,733	0.52
2001	3,650	0.94
2002	2,448	0.61
2003	2,191	0.52
2004	3,572	0.75
2005	6,846	1.17
2006	11,867	1.66
2007	11,000	1.31
2008	26,450	2.66
2009	26,935	2.24
2010	26,000	2.03
2011	30,000	2.14
2012	34,000	2.13

According to above Table 05, Government expenditure of fertilizer subsidy has increased over time while comparative reduction after year 2007 due to identification of inefficiency of fertilizer subsidy programme. Market price of main subsidized fertilizer has increased over time in Sri Lanka leading high cost to the government to allocate subsidy. Since 2009, there was an obligation to get eligibility for fertilizer subsidy. To be eligible for the fertilizer subsidy, farmers have been required to sell 500 kg of paddy per hectare to the government at a guaranteed price. This policy has not been financially attractive to farmers because the market price is usually higher than the guaranteed price. (Weerahewa et al, 2010).

Cost of production of Paddy

Concerning cost of production and profitability of paddy farming in Sri Lanka, current cost of production of rough rice is Rs. 8.57 per kg. The cost of labor, farm power and tradable inputs constitutes 55%, 23% and 23% respectively. The labor cost has risen at a higher rate than other costs over the last few years (Department of Agriculture, 2013). Based on the literature analysis, cost incurred by fertilizer application contrast to other cost components of paddy cultivation is nearly 6 percent of total cost (Prasanna et al, 2011). Further, difference of total cost of production of fertilizer subsidy and without fertilizer subsidy is only about 20 percent (Department of Agriculture).

Issues in subsidy programme in Sri Lanka

Paddy Farmers in Sri Lanka will eligible to apply for fertilizer subsidy, if they have legal title to their paddy lands. The subsidized fertilizer distributes through Agrarian Service Centers (ASC) under Agrarian Services Department. Before every cultivation season, the ASCs call for applications from eligible farmers, who are required to furnish information on which crops they cultivate, the amount of land devoted to each crop, and the amount and type of fertilizer required. Finally, qualified farmers have to visit ASCs to collect the fertilizer pack. Among the very recent drawbacks of the fertilizer subsidy programme, some major issues have identified by referring previous research study on Fertilizer subsidy programme in Sri Lanka. Farmers use less fertilizer per acre than recommended since farmers cultivate more land than they legally own land by illegal land encroachment. Therefore, expected outcome of the fertilizer subsidy does not happen in some context. Further, Informal sales of fertilizer between well-off and worse-off farmers may lead to overuse of fertilizer by well-off farmers and underuse of fertilizer by worse-off farmers. Inefficiencies of the current distribution system have created opportunities for some dishonest government officials to get black money.

However, a study by farmers are willing to pay about US$9.23–11.54 (Rs.1,000–1,250) per 50-kg bag if fertilizer can be made readily available in the local open market, compared with the current subsidized rate of US$3.23 (Rs. 350) Wijetunga, et al (2008) as cited by Weerahewa et al (2010). Farmers' willingness to pay is high due to transaction cost incurred by farmers, who miss about three working days to obtain the fertilizer and transportation costs. This indicates the necessity of having a farmer-friendly fertilizer distribution system.

Since last decades, there is an emerging public and scientific debate on health issues associated with Kidney diseases in major paddy growing areas in Sri Lanka. Annually, many farmers affected by kidney diseases due to chemical affected water consumption. At present, fertilizer subsidy allocates to all farmers without having target mechanism and this has cause overuse of fertilizer. This has led to nutrient imbalances in the soil causing long-term adverse impact on soil fertility and improper and excessive fertilizer application, particularly nitrates and phosphates have adverse impacts on human health and on the environment.

5. Conclusion

Generally, the fertilizer subsidy programmes in developing countries criticize as ineffective programme compare to the cost incurred by particular government. Especially misuse and corruption are happening during distribution process. This phenomenon has common to Sri Lankan context. Since the government expenditure has risen during last two decades for the fertilizer subsidy, it has created considerable social impact. By analyzing above study findings and literature based it can be generally concluded that fertilizer usage can be influenced by fertilizer price, producer price of paddy and extent under paddy cultivation. In generally, it can conclude that fertilizer subsidy has been achieved rice self-sufficiency and increased production is therefore success in terms of production scenario. However, this self-sufficiency may be affected by other technological and farm management and crop management practices rather than subsidy programmes.

Paddy market and fertilizer market has affected by fertilizer subsidy programme. Cost of production of paddy has not greatly affected due to fertilizer subsidy. Anyhow, government expenditure is still growing for allocation of subsidy scheme. Following are the few of suggestions to improve fertilizer subsidy programme in Sri Lanka based on study findings and literature review. Introduce alternative instrument for present subsidy programme such as voucher scheme, which is practicing in Malawi and some other African countries, seems to be good alternative to Sri Lanka. Meanwhile, rearrange the subsidy programme giving priorities based on important characteristic such as target group, agro ecological region, poor region of country, climatic condition and crop varieties etc. Present Fertilizer subsidy distribution process is lengthier and inconvenient to farmers causing high transaction cost. Finally, fertilizer subsidy programme in Sri Lanka has concluded as economically ineffective, but politically and socially acceptable programme.

6 References

Afua, B. B. (2010) Old Problems in the New Solutions? Politically Motivated Allocation of Program Benefits and the "New" Fertilizer Subsidies. *IFPRI Discussion Paper 01002*, International Food Policy Research Institute, P. 2-3.

Andrew D and Chirwa, E. (2011) The Malawi Agricultural Input Subsidy Programme: 2005-6 to 2008-9, Forthcoming (after further editing*), International Journal of Agricultural Sustainability* (IJAS), 9(1), p. 10-15.

Central Bank of Sri Lanka. Various years, *Annual report*, Colombo, Sri Lanka.

Chandrasiri, W.A.C. K and Karunagoda K.S. (2008), Technical efficiency of paddy production in the North and North Western Provinces of Sri Lanka. *Proceeding of the Second Annual Research Forum of the Sri Lanka*, Agricultural Economics Association, October 3.

Chibwana, C and Fisher, M. (2011), The impacts of agricultural input subsidies in Malawi. International Food Policy Research Institute, Policy Note 5 / March 2011.

Department of Census and Statistics. Various years. *Statistical abstract*. Colombo, Sri Lanka.

Department of Agriculture, various years, *Annual Publication and Research study*. Colombo, Sri Lanka.

Ekanayake, H.K.J. (2006), Impact of fertilizer subsidy in paddy cultivation in Sri Lanka, *Staff Studies*, Central Bank of Sri Lanka, p 73–92.

Ellis, F. (1992), Agricultural Policies in Developing Countries. Cambridge University Press, Cambridge, New York, p, 125-150.

Ephraim Chirwa, et al. (2011), Conceptualizing Graduation from Agricultural Input Subsidies in Malawi. *Working Paper* 029, p 02.

Takeshima, H and Haklim L, (n.d) Agricultural Inputs Subsidy and Their Development Impact. Conventional Wisdom, *Policy note 01,* International food policy research Institute.

Institute of Development Studies. (2008), Rethinking Agricultural Input Subsidies in Poor Rural Economies, Future Agricultures. Institute of Development Studies, Brighton.

Institute of Policy Studies. (1992), The Sri Lanka Economy: On a Path of Progress. Colombo, Sri Lanka.

Gilbert, J.R and Jayne, T.S. (2012), Do Fertilizer Subsidies Boost Staple Crop Production and Reduce Poverty across the Distribution of smallholders in Africa? Quintile Regression Results from Malawi. International Association of Agricultural Economists (IAAE) Triennial Conference, Foz do Iguaçu, Brazil, 18-24 August 2012.

Jeminiwa A, Charles. (2011), Towards Improved Fertilizer Subsidy Programme in Nigeria: Drawing Lessons from Promising Practices in the Sub Saharan Africa. Report Prepared for the Visiting Fellowship for African Policy Officials, October, 2011.

Masao, K and Aluwihare, P.B, (1990), Fertilizer Response Functions of Rice in Sri Lanka: Estimation and some application. International Irrigation Management Institute, Colombo, Sri Lanka.

National Fertilizer Secretariat. Various years. *Annual reports*. Colombo, Sri Lanka.

Nicholas M and Todd B. (2009), Fertilizer Subsidies in Africa, Are Vouchers the-Answer? International Food Policy Research Institute, *IFPRI Issue Brief 60*, July 2009.

Nurul NR, et al. (2012), The impact of fertilizer subsidy on Malaysia Rice/paddy Industry using a System Dynamic Approach. *International Journal of Social Science and Humanity*, 2(3).

Pranab B. and Mookherjee, D. (2011), Subsidized Farm Input Programmes and Agricultural Performances: A farm level Analysis of West Bengal´s Green Revolution 1982-1995. *Seminar paper, Revised Version*: February 8, 2011. p 3-4.

Prasanna, R.P.I.R et al. (2011), Factors Affecting Farmers` Higher Gain from Paddy Marketing: A Case Study on Paddy Farmers in North Central Province, Sri Lanka. *International Journal of Agricultural Management & Development (IJAMAD)*. p 57-69.

Rajapaksha, R.D.D.P, and Karunagoda, K.S. (2008). Fertilizer demand for paddy cultivation in Sri Lanka: A profit function approach. *Proceeding of the Second Annual Research Forum of the Sri Lanka*, Agricultural Economics Association.

Samaratunga, P. et al. (2006), Regional Agricultural Trade Liberalization Efforts in South Asia: Retrospect and Prospects. Asia-Pacific Research and Training Network on Trade, *Brief No. 10*, December 2006.

Saweda, L.T et al. (2010), A Review of Fertilizer Policy Issues in Nigeria. NSSP Working Paper No. 0019, Nigeria Strategy Support Program (NSSP), International Food Policy Research Institute, Nigeria, p 16.

Steven R. Tabor and Samaratunga, RHS. (1995), Agricultural Research in an Era of Adjustment: Policies, institutions and Progre*ss*. The International Bank for Reconstruction and Development, The World Bank, Washington, D.C.

Timmer, P.C et al. (1983), Food Policy Analysis*.* World Bank. Washington, D.C.

Weerahewa, J. et al. (2010), The Fertilizer Subsidy Program in Sri Lanka, A case study of the programme: Food Policy for Developing countries: the Role of Government in the Global food system. International Food Policy Research Institute.

Weerahewa, J. (2004), Impacts of trade liberalization and market reforms on the paddy/rice sector in Sri Lanka. Markets, Trade and Institutions Division Discussion Paper No. 70. Washington, DC: International Food Policy Research Institute. Available from*: http://www.ifpri.org/divs/mtid/dp/papers/mti_dp70.pdf. [Accessed 12th March 2013]*

Wijetunga, WMLK et al. (2008). Impact of "KETHATA ARUNA" fertilizer subsidy on paddy production in Minipe Scheme. *Proceeding of the Second Annual Research Forum of the Sri Lanka Agricultural Economics Association*, October 3.

World Bank. (2003), Sri Lanka Promoting Agricultural and Rural Non-farm Sector Growth. *Volume 1: Main Report, Rural Development Unit, South Asia Region*, World Bank, February 26, 2003.

The importance of institutions on export performance:
The case of Kosovo

Linda Tahiri Rukiqi[1]

[1]South-East European Research Centre (SEERC),
International Faculty, The University of Sheffield,
ltahiri@seerc.org

Abstract. Despite Kosovo's robust economic growth in the past decade and a half, the country remains the poorest in Europe and its deep trade deficit is a persistent problem. In order to boost development and facilitate growth, pragmatic policies and export incentives are crucial. This short paper emphasizes the importance of exports for the development of economies and the significance of quality of institutions on the volume of exports. In order to estimate the impact of institutions on the export volume of Kosovo, the short paper proposes to conduct a survey of businesses to find answers on how businesses view export policies, what the main constraints are, and what incentives would be helpful, and conduct in-depth interviews of businesses and main stakeholders to explore key aspects in more detail.

Keywords: *formal institutions, informal institutions, exports, trade balance, Kosovo*

1.   Introduction

The positive effects of international trade on economic performance have been widely acknowledged by economists, starting with Smith (1776). According to Krugman (1981), a pioneer of new trade theory, for example, international trade permits cost reduction because of economies of scale; therefore there are potential gains to all trading nations. Van den Berg and Lewer (2007, p. 76) explain how "a sudden shift in trade policy that opens up new trade provides an immediate gain in real per capita income, which, in turn, accelerates technological progress and increases the rate of economic growth permanently". While substantive literature has stressed on the positive effects of international trade on economic performance, a parallel literature has documented the importance of high institutional quality for economic growth and development (North, 1990; Acemoglu et al., 2001; Levchenko, 2007), indicating that quality institutions are a

prerequisite for trade to impact growth. Both formal institutions, or formal rules and regulations enforced through official channels, and informal institutions, such as social norms and codes of conduct, can impact a country's trade performance.

Kosovo, a small economy in transition, has made significant progress with building formal institutions and has experienced robust economic growth after its liberation from Serbia in 1999; however, its GDP/capita remains the smallest in Europe and the deep trade deficit continues to be a persistent problem. To improve the performance of the economy and facilitate further growth, pragmatic economic policies and incentives to increase exports are considered crucial by both international institutions, such as IMF and the World Bank, and local experts. Given the importance of trade in economic development, and the belief that quality institutions are a prerequisite for trade to impact the performance of economies, Kosovo needs to have the institutional framework which promotes exports.

This short paper emphasizes the importance of exports on development of economies and the significance of quality of institutions on the volume of exports. The following section reviews the existing literature. Section 3 presents the proposed methodology to estimate the impact of institutions on the export volume of Kosovo, whereas section 4 concludes.

2.   Literature review

The positive effects of international trade on economic performance have been studied by numerous economists, starting with Smith (1776). Different theories aim to explain international trade, most of them focusing on the explanations of trade flows between at least two nations, the benefits and losses of trade across borders, and the effects of trade policies on economic development. Krugman (1981), a pioneer of new trade theory, explained that because of economies of scale, cost reduction is permitted on a global scale therefore there are potential gains to all trading nations. While two countries may have no noticeable differences in opportunity costs at a particular point in time, specialisation in a particular industry by a country gives it benefits, because specialization increases output thus they gain from economies of scale.

Understanding the relationship between trade openness and the development of economies has very important implications for policy makers, hence numerous empirical studies have been conducted to estimate the effects of trade on economic growth (Maizel, 1963; Kavoussi, 1984; Rodriguez and Rodrik, 2000; Dollar and Kraay, 2003). Kavoussi (1984) has examined the link between trade openness and economic growth in 73 middle and low income developing countries, and found that higher rates of growth are strongly correlated with a higher level of exports for both groups of countries, middle and low income. In the same line, a study of Maizel (1963) which investigated the impact of international trade on economic growths in developed countries has found that there is a

strong positive correlation. Sachs and Warner (1995) constructed an indicator of openness classifying an economy as closed or open based on tariff rates, non-tariff trade barriers, the size of the black market, whether or not it had a socialist system, and whether or not there was a state of monopoly of major exports. They found that the average economic growth rate in open economies is significantly higher in comparison to the growth rate in closed ones. According to them, all countries that liberalized trade show economic growth.

While the positive effects of international trade on economic performance have been widely acknowledged by economists, a parallel literature has stressed on the importance of high institutional quality for economic growth and development (North, 1990; Acemoglu et al., 2001; Levchenko, 2007), indicating that quality institutions are determinants of trade performance, hence they are essential for trade to impact growth and development. North (1990, p. 3) defines institutions "as rules of the game in a society or, more formally, [they] are constraints that shape human interaction". Academic debate has come to employ the institutional lens in different ways, which has led to different interpretations and classifications, but they are primarily categorized into formal and informal institutions. Formal institutions are comprised of rules and procedures generated and enforced through official channels (Ellickson, 1991; Knight, 1992; Calvert, 1995), such as state agencies including courts and legislatures, state-enforced rules comprising of constitutions, laws and other regulations, as well as rules that govern corporations, political parties and other non-state organizations. Whereas, informal institutions are rules and procedures generated and enforced outside the officially sanctioned channels (North, 1990; O'Donnell, 1996), including social norms, customs, and traditions. According to Williamson (2000) informal institutions define the basic foundations for a society's institutions and they influence formal rules. Furthermore, Roland (2004) notes that these institutions change more slowly compared to other institutions, therefore, have an important effect on the choice of formal institutional environment comprised of political, economic, and legal institutions. Transition economies often reform formal institutions to resemble those of developed economies; however, given that informal institutions define the basic foundation for a society's institution, transition economies cannot reap the same benefits more developed countries do.

Both formal and informal institutions lay behind the reason for differences in growth rates and development paths across developing and developed countries. Acemoglu, Johnson and Robinson (2001), for example, state that institutions matter for economic growth because they shape the incentives of key economic actors in society. Rodrik (2007, p.154) argues markets "are not self-creating, self-regulating, self-stabilizing, or self-legitimizing", therefore they require being backed up by a wide range on institutions."; since well-functioning markets cannot be developed overnight, rules and regulations, including competition laws, public procurement laws and taxes, are a necessity for markets to grow.

Although Kosovo has made significant progress with building formal institutions and has experienced robust economic growth in the past decade and a half, it's GDP/capita remains the smallest in Europe, and the trade deficit and unemployment continue to be persistent problems. The table below shows selected economic indicators of year 2013.

Table 1.: Economic indicators, 2013

GDP (mil. €)	GDP growth	GDP/capita (€)	Unemployment	Imports (mil. €)	Exports (mil. €)	FDI inflow (mil. €)
5,125[1]	3.2%[1]	2,757[1]	30.9%[2*]	2,306[1]	305[1]	258,5[3]

Source: [1]Ministry of Finance of the Republic of Kosovo (2014), [2]Kosovo Agency of Statistics (2013), [3]Central Bank of the Republic of Kosovo (2014).

*Data for 2012

In order to improve the performance of economy, pragmatic economic policies and export incentives are required. To be able to recommend further actions, the institutional framework for export should be examined, export policies should be viewed, main constraints for exporting should be defined, and incentives that lead to growth of export volume should be identified.

3.   Proposed methodology

In order to estimate the impact of institutions on the export volume of Kosovo, the finding should focus on both a survey and in-depth interviews. First, a survey of businesses should be conducted. Following, in-depth interviews of stakeholders and in-depth interviews of businesses should be performed.

To conduct the survey of Kosovo businesses, an updated list of exporting businesses that export over 1 million Euros should be obtained from the Kosovo Customs. The survey should include questions that aim to find answers on how businesses view export policies, what the main constraints are, and what incentives would be helpful. On the other hand, the short paper proposes to perform in-depth interviews with main stakeholders, including the Ministry of Trade and Industry, Chambers of Commerce, and relevant Agencies. These interviews are crucial to examine the institutions framework in Kosovo. Furthermore, in-depth interviews should be performed with a part of businesses,

selected on basis of sectors and importing groups of countries, in order to explore key aspects in more detail.

4.   Conclusion

Despite Kosovo's robust economic growth in the past decade and a half, the country remains the poorest in Europe with deep trade deficit as a persistent problem. In this line, pragmatic policies and export incentives are crucial. This short paper emphasizes the importance of exports on development of economies and the significance of quality of institutions on the volume of exports.

In order to estimate the impact of institutions on the export volume of Kosovo, the short paper proposes to conduct a survey of businesses to see how businesses view export policies, what the main constraints are, and what incentives would be helpful, and to conduct in-depth interviews of businesses and main stakeholders to examine the institutional framework in Kosovo and explore key aspects in more detail.

References

Acemoglu, D. et al, (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, 91 (5), p. 1369-1401.
Acemoglu, D. and Robinson, J., (2012). *Why nations fail: the origins of power, prosperity, and poverty,* Crown Publishers, New York.

Calvert, R. (1995). The Rational Choice Theory of Social Institutions: Cooperation, Coordination, and Communication.

Central Bank of the Republic of Kosovo (2014). *Monthly Statistics Bulletin, April 2014*, Prishtina.

Dollar, D. and Kraay, A. (2003). Institutions, trade, and growth. *Journal of Monetary Economics, Elsevier*, 50 (1), p. 133-162

Ellickson, R. (1991). Order Without Law: How Neighbors Settle Disputes. Cambridge, MA: Harvard University Press.

Kavoussi, R. (1984). Export Expansion and Economic Growth: Further Empirical Evidence. *Journal of Development Economics* 14, p. 241-250.

Keefer, P. and Knack, S., (1995). Institutions and Economic Performance: Cross-Country Test Using Alternative Institutional Measures. *Economics and Politics*, 7 (3), p. 207-227.

Knight, J. (1992). Institutions and Social Conflict. Cambridge, UK: Cambridge University Press.

Kosovo Agency of Statistics (2013). *Labour Force Survey 2012.*, Prishtina.

Krugman, P. (1981). Trade, Accumulation, and Uneven Development. *Journal of Development Economics* 8: 149-161.

Levchenko, A. (2007). Institutional Quality and International Trade. *Review of Economic Studies* 74, p. 791-819.

Maizel, A. (1963). Industrial Growth and World Trade. Cambridge, the University Press.

Ministry of Finance of the Republic of Kosovo (2014). *Medium Term Expenditure Framework 2015-2017*, Prishtina.

North, D., (1990). *Institutions, Institutional Change, and Economic Performance,* Cambridge University Press, New York.

O'Donnell, G. (1996). Illusions about Consolidation. *Journal of Democracy*, 7(2), p. 34-51.

Rodriguez, F. and Rodrik, D. (2000). Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence, *NBER Macroeconomics Annual* 15, p. 261–325

Rodrik, D. (2007). *One Economics, Many Recipes: Globalization, Institutions, and Economic Growth.* Princeton, NJ: Princeton University Press.

Sachs, J., and Warner, A., (1995). Economic Convergence and Economic Policies. *Brookings Papers in Economic Activity*, 1 (1), p. 1–95.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations,* 1. London: W. Strahan.

Van den Berg, H. and Lewer, J. (2007). *International Trade and Economic Growth*, Armonk, NY: M.E. Sharpe.

Williamson, J., (2000). What Should the World Bank Think About the Washington Consensus?, *World Bank Research Observer,* 15 (2).

The Implementation of Fiscal Policies in the Balkan Region and Their Impact on Economic Growth- Lessons to be learned

Ardian Syla[1]

[1] South East European Research Centre (SEERC),
CITY College – The University of Sheffield International Faculty
adsyla@seerc.org

Abstract. Given the importance of economic stability and considering that economic growth is very crucial for every country, it is important to understand which policy tools are more effective for the government. This paper aims at explaining fiscal policy (government spending and increase in taxation) and its effects on economic growth for the Balkan Region. The literature review ascertains that fiscal policy has mixed effects on the economic growth of a country. However, most of the literature findings focus on how a well managed fiscal policy can lead to economic growth. Using data from 11 countries from the Balkan Region for the period 2000-2016 we will be testing the correlation between fiscal policies and economic growth. The empirical evidence obtained from this study may enrich our knowledge on the complex relationship between fiscal policies and economic growth for the Balkan Region.

Keywords: *Fiscal Policy, Economic growth (GDP), Balkan Region, Correlation*

1   Introduction

Economic growth is an imperative objective of developed and developing nations. In order to attain this objective, certain paths and policies are needed. Growth policies are essential tools to accelerate the pace of economic growth.

Mises (2006) explains that in growth policies, there are no miracles. Fiscal policy involves government changes at the level of taxation and government spending with the aim to influence aggregate demand in order to have an impact on the level of economic activity (Mishkin, 2007a). The main purpose of fiscal policy is to reduce the rate of inflation, stimulate economic growth in a period of recession and in general terms to stabilize economic growth. There are two economic schools, which argue differently with regards to fiscal policy. The Keynesian school argues that fiscal policy is a powerful tool which increases output and employment when the economy is operating below its national output (Von Hagen and J-Mundschenk, S.2003). On

the other hand, monetarists believe that government spending and tax change can have only temporary effect on aggregate demand, output and employment, (Viren, M.1998).

The main philosophy of the fiscal policy is based on the theory which states that government actions can influence macroeconomic productivity levels by increasing or decreasing tax level and public spending. The Keynesian theory was one of the most fundamental achievements which reoriented the way economics viewed the influence of government activities on the economy. Before Keynes, it was believed that government spending and taxation were powerless to affect aggregate levels of spending and employment in the economy.

This study will apply the existing econometrical models on countries (Balkan Region) which have not yet been explored. After reviewing existing literature on fiscal policies and their impact on economic growth, we easily come to conclusion that there are no existing studies regarding fiscal policies and their impact on economic growth for the Balkan Region as a whole. The possible outcomes from this study will be a benefit for the literature, future studies and policy makers.

2   Literature Review

2.1   Fiscal Policy – Government Spending

There is much literature on the impact of monetary policy and its effect on economic growth, however, fiscal policy has received less attention and was believed to have an irrelevant effect on economic development.

Nevertheless, in contemporary literature this subject is receiving much greater attention from academia, central banks and governments. It is generally believed that fiscal policy is associated with growth, or more precisely, it is held that suitable fiscal measures in specific cases can be a strong stimulator for economic development or economic growth (Barro, 1991, Barro and Sala-i-Martin,1995). Government expenditures can have an affirmative impact on the economic growth through increasing the production, and thus having direct effect in output growth (Baron, 1976).

However, it is important to remember that public expenditures such as investment in public infrastructure have retreating marginal return, and thus government spending might become inefficient (Fischer, 1989). Furthermore, Tanzi and Schuknecht (1997) assessed the incremental impact of public spending on social and economic indicators such as real growth. Asghar et al. (2011) in their study point out that the resources allocated by the government to education and health sector contribute to economic growth.

Findings from empirical evidence which explain the linkage between public expenditures and economic growth are to some degree mixed. In most cases, the literature uncovers an inverse relationship between government spending and economic

growth (Fischer, 1997, Barro and Martin, 1991; Devarajan et al., 1996). However, it seems that this relationship is positive between the increase in public expenditures and the economic growth or growth rate (Mishkin, 1982, Barro, 1999). According to Martina Dalić (2013), fiscal policy can have positive effects on economic growth by intervening in the structure of total expenditures such as reductions on current expenditures, lower taxes and higher government investments.

2.2   Fiscal Policy – Interest Rates

On the other hand, fiscal policy has to do also with increase or decrease in interest rates. Gale and Orszag (2004) argue that there are two reasons to explain how budget deficit raises nominal interest rates; the first one is that public deficit causes a reduction on aggregate savings if the private savings do not increase by the same amount, and considering that there is no compensation from foreign capital inflows, this would lead to a decrease in the supply of capital. The second reason is that deficits increase the government debt and as a result the outstanding amount of government bonds. However, some studies find out that interest rates are likely to increase after a rise in the deficit, whilst others do not (Engen and Hubbard, 2004).

Fiscal expansion means government spending, and spending means higher budget deficits. As government debt increases by practicing fiscal expansion   an increase in inflation risk reinforces crowding out effects through interest rates (Miller, Weller, and Skidelsky 1990). Under these conditions, a onetime fiscal expansion will be more efficient than a permanent one. Blanchard and Summers (1984), came to conclusion that most increases in interest rates during 1970s are attributed to higher stock market returns. However, Barro (1992) concludes that even a one percent of GDP increase in debt leads to an increase of interest rates by 12 basis points. Barro and Salai Martin (1995) cannot conclude in their study about the significance effect of either deficits or debt on interest rates, while Tanzi and Lutz (1993) support that either debt or deficit, but not both, have an important role in determining interest rates.

2.3   Tax Increase versus Government Spending

Implementing fiscal policy in a country is just a matter of decision; however, in many cases fiscal policy should be considered as the primary choice in economy if compared to monetary policy which stimulates the economy in the short run. In the literature, the subject is largely explored and discussed. Giavazzi and Pagano (1990) are the very first authors to disagree that large deficit reduction on the spending side could be expansionary. In theory, expansionary impacts of the fiscal consolidation go through both demand and supply side (Strauch, 2001). Decrease in interest rates leads to appreciation of bonds and stocks while g enerating wealth f o r investors (Blanchard, 2002).

On the other hand, the supply side, expansionary effects of fiscal consolidation goes through the labor market and if taxes are increased and spending is cut, they have a direct effect on the individual labor supply, which then can create an

imperfectly competitive labor market (Alesina and Ardagna, 2004). According to Faini (2005), the impact of fiscal policy depends on the key parameters, such as how Ricardian the economy is. Both a tax cut and an increase in public spending will boost aggregate demand, which later on shifts the investment curve to the right, wages increase, employment increases, prices go up and as a result there is a raise in the short run interest rates (Mishkin, 2007b).

2.4   The Duration of Fiscal Consolidations

When fiscal policy is implemented it takes time to consolidate and have an effect on economic growth. Wenzelburger and Wagschal (2008) focused their research on 23 OECD countries from the period 1945-2000 and observed 26 cases of fiscal consolidation. Further with their studies, they found out that from these 26 cases, 15 were successful and 10 were unsuccessful and the average consolidation period lasted about 3.1 years. Another study conducted by Zaghini (1999), observed 98 consolidation events in 14 European countries from 1970 to 1998 and the main finding of this study is that on average successful consolidations lasted about 2.7 years. Guichard et al., (2007) monitored 85 fiscal adjustments in 24 countries in the period of 1978-2005 and concluded that the median duration of these fiscal adjustments was two years. According to these findings the longest consolidation were Sweden which lasted from 1981-87 and 1994-2000, and eight years in Japan (from 1980-87).

3   Methodology

The methodology that will be applied in conducting this study with regards to fiscal policy and economic growth for the Balkan Region (Albania, Greece, FYROM, Serbia, Bosnia and Herzegovina, Croatia, Romania, Bulgaria, Kosovo, Montenegro, Slovenia) will be performed in several steps and different research techniques will be employed such as The Engle and Granger (1987) two-step technique, VAR, OLS, Panel Analysis, One way and two way Anova.

3.1   Engle and Granger –two-step technique

The Engle and Granger Causality two-step technique is a very important testing method since it can determine the integration and co-integration of both the series (X_t and Y_t). Variable Y which in our case stands for economic growth is assumed to be caused by X, which in our case stands for fiscal policies.

3.2   Vector Autoregression (VAR)

The vector autoregression (VAR) model is one of the most applied and successful tool for analysis regarding multivariate time series data. This technique consists of the extension of the univariate autoregressive model to multivariate time series. It

is a proven tool, flexible and very useful tool for describing and forecasting the behavior of economic and financial time series. Also, the VAR model is also applied for the structural inference and policy analysis.

3.3 Ordinary Least Square – OLS

In addition, we apply OLS and panel econometric techniques. OLS assume that the error in each time period is uncorrelated with the explanatory variables in the same period. It allows for more efficient parameter estimates, uncovers dynamic relations and identifies otherwise unidentified models. So, we initially estimate our models by OLS and select the appropriate model specification using the Akaike Information and Schwartz Bayesian Information Criteria as selection criteria. In light of that, my research assumes that there is a time-constant unobserved effect, which may represent country-specific technology, tastes, historical and cultural factors and proceed with fixed effects estimation.

4 Conclusions

This study aims to investigate the eventual relationship between fiscal policies and economic growth in the Balkan Region (Albania, Greece, FYROM, Serbia, Bosnia and Herzegovina, Croatia, Romania, Bulgaria, Kosovo, Montenegro, and Slovenia) for the period 2000-2016. Most of the studies suggest that there is a relationship between fiscal policies and economic growth. By employing the above mentioned methodologies this study will empirically investigate the eventual relationship between these two variables. The results from most of the studies regarding fiscal policies and their impact on the economic growth show that fiscal policies play an important role toward the economic growth, and therefore we expect the similar relationship from this study.

5 References

Alesina A. and R. Perotti (1998), "Fiscal Adjustment in OECD Countries: Composition and Macroeconomic Effects", IMF Staff Papers, 44 (2)

Asghar, N. et al (2011), "The Impact of Government Spending on Poverty Reduction: Evidence from Pakistan 1972 to 2008.

Ardagna S., F. Caselli, T. Lane (2004), "Fiscal Discipline and the Cost of Public Debt Service: Some Estimates for OECD Countries." NBER Working Paper No 10788 .

Robert J. Barro (1999), " Inequality, Growth, and Investment", NBER Working Papers 7038, National Bureau of Economic Researches, Inc.

Barro, Robert (1991), "Economic Growth in a Cross Section Countries", The Quarterly Journal of Economics. Vol 106, No. 2, pp. 407-443

Barro, Robert, Sala-i-Martin, Xavier (1991), Convergence across States and Regions, Harvard University

Barro, Robert, and Sala-i-Martin, Xavier, (1995), Economic Growth. New York: McGraw-Hill.

Barro, Robert (1992), " Public Finance in Models of Economic Growth", Review of Economic Studies, Wiley Blackwell

Blanchard, Olivier and Roberto Perotti (2002), "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output". Quarterly Journal of Economics.

Blanchard, Oliver and Summers Lawrence (1984), " Perspectives on High World Real Interest Rates, Massachusetts Institute of Technology

Devarajan Sh. et al (1996), " The Composition of Public Expenditures and Economic Growth". Policy Research Department, The World Bank, Washington DC

Dickey, D.A. and Fuller, W.A. (1981), "Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root. Econometria, 49, 1057-1072.

Engle, R.F. and Granger, C.W.J. (1987),"Cointegration and Error Correction: Representation, Estimation and Testing. Econometrica, 55, 251-276.

Engen E. and G, Hubbard (2004), "Federal Government Debts and Interest Rates." NBER Working Paper n. 10681, NBER, Cambridge.

FAINI, R. (2005). Fiscal policy and interest rates in Europe. PMEP Working Paper, Luxembourg: 41st Panel meeting of Economic Policy

Fischer, S., Dornbusch, R. (1997) Macroeconomia, Editura Sedona, Timişoara,.

Gale, W. and P. Orszag (2004), "Budget Deficit, National Savings and Interest Rates", Brookings Papers on Economic Activity, 2, 101-210.

Gale W. and P. Orszag (2002), "The Economic Effects of Long Term Fiscal Discipline." Urban Brookings Tax Policy Center Discussion Paper

Giavazzi, Francesco, and Marco Pagano(1990), "Can Severe Fiscal Contractions be Expansionary?

Guichard, S., Kennedy, M., Wurzel, E., and Andre, C. (2007), "What promotes Fiscal Consolidation." OECD country experiences, OECD Economics Department Working Papers No. 553.

Hagen, Jiirgen, Andrew Hughes-Hallett, and Rolf Strauch (2003), "Budgetary Consolidation in EMU," CEPR Discussion Paper No. 148 (London: Centre f or Economic Policy Research).

Jaimovich, Nail and Sergio Rebelo (2008):"Can News about the Future Drive the Busi- ness Cycle?", American Economic Review forthcoming.

Miller, Marcus, Robert Skidelsky, and Paul Weller (1990), "Fear of Deficit Financing—Is it Rational?" in Public Debt Management: Theory and History, ed. by Rudiger Dornbusch and Mario Draghi (Cambridge and New York: Cambridge University Press), pp. 293-310.

Mises L.M. (2006). Economic policy, Thoughts for today and tomorrow, Ludwig Von Institute

Mishkin, F.S. (1982),"Does Anticipated Monetary Policy Matter." An econometric investigation. J. Politic. Econ., 90: 22-51.

Martina D. (2013), "Fiscal Policy and Growth in New Member States of the EU: a Panel Data Analysis. Financial 4/2013 Theory& Practices

Mishkin, Frederic S. (2007). "Housing and the Monetary Transmission Mechanism," Finance and Economics Discussion Series 2007-40. Washington: Board of Governors of the Federal Reserve System, September, www.federalreserve.gov/pubs/feds.

Afonso, A.; Schuknecht, L.; and Tanzi, V. (2005) "Public Sector Efficiency: An International Comparison." Public Choice 123: 321–47.

Viren, M. (1998), "Do the OECD Countries Follow the Same Fiscal Policy Rule? Government Institute for Economic Research Discussion Paper 187

VON HAGEN J.-MUNDSCHENK S. (2003), "The Functioning of Economic Policy Coordination, ZEI Working Paper, B 08.

TANZI V. and LUTZ M., 1993, Interest rates and government debt: are the linkages global rather than national?, in: VERBON H. and VAN WINDEN F., The Political Economy of Government Debt, Amsterdam: Elsevier Science, 233-253.

Wagschal, U. and Wenzelburger, G. (2008), "Successful Budget Consolidation." An international Comparison.

Zaghini, Andrea (2000), "Increasing Public Expenditures: Wagner's Law in OECD Countries. European Central Bank

Gender differences in building brand loyalty through Social Media: the case of Balkans

Zlatina Gubatova and Alexandros Kapoulas

Business Administration and Economics Department
City College – International Faculty of the University of Sheffield
zlgubatova@city.academic.gr, akapoulas@city.academic.gr

Abstract. The report aims at investigating academic literature regarding gender differences in building brand loyalty through Social Media, as a basis for further investigation in the Balkan region. By utilizing previous findings, the authors attempt to uncover main aspects that affect customer brand loyalty in the context of Social Media. Thus, there are four main points discussed throughout the report. First, the report draws upon the notion of brand loyalty as a means through which companies can differentiate themselves by going well beyond physical characteristics to appeal to customers on an emotional level. Second, previous research regarding the key role of brands as a self-expression tool is employed. Third, concepts such as "social mobility", "satisfaction with virtual brand communities", "brand love", and "trust", are presented within the context of brand communities as a tool for shaping brand loyalty. Finally, discussion about virtual brand communities, eWOM, and gender differences in building e-loyalty as predictors of brand loyalty takes place. The report does not contain empirical analysis, as it presents a PhD proposal. The research adopts qualitative methodology based on the reasoning that it will be conducted for a first time in the Balkan region as well as on the opportunity that qualitative research allows to the investigator to build closer relationships with the participants, which will further lead to acquiring more honest responses.

Keywords: *social media, gender, differences, Balkans, brands, loyalty*

Introduction

The advancement of Social Media has allowed companies to place themselves closer to customers, establish communication on a constant basis, and engage customers on an emotional level. Moreover, the development of this medium has raised issues such as adopting innovative ways through which companies can build loyal customer base-regardless of competitors' actions (Aaker, 2012; Tiwary and Roy, 2012). Thus, via comprehending and taking advantage of Social Media, obtaining thorough insights into the behavior of Social Media consumers, and gender differences among customer

groups, companies can benefit from creating a loyal customer base, regardless of the location their operations take place.

In this context, the Balkan region presents an attractive area for further investigation, as current academic literature does not provide data in this matter. Hence, it is suggested that findings will allow to fill gaps in academic literature as well as to present valuable guidelines for companies, which conduct their business in this area.

Literature review

Brand Loyalty

According to Jung and Shen (2011) brands can differentiate themselves within a product category not merely by physical characteristics, but by emphasizing on intangible (symbolic) benefits for building loyal customer base (Power and Hauge, 2008), which presents the ultimate path towards companies' success (Tsiotsou, 2013). In this context, social identity is useful for explaining consumer behavior (Chan *et al.* 2013) from the perspective of brand loyalty for maintaining customer-brand relationships, regardless of competitors' actions (Aaker, 2012; Tiwary and Roy, 2012) Brand literature accepts two main factors to shape brand loyalty: behavioral loyalty, standing for customer's purchase of a brand over time, and attitudinal loyalty, characterizing customers' affiliation towards a certain brand, impacting repurchase behavior (Chahal and Bala, 2010, cited in Smith, 2012; Gecti and Zengin, 2013; Abraheem and Mohammad, 2012).

Regarding the latter, Yousaf *et al.* (2012) recognizes the role of brand loyalty as a moderator for increased sales and profit margin, as emotional attachment to a brand (Rossiter and Bellman, 2012), customer engagement (Vivek *et al.* 2012) and customer satisfaction (Voss *et al.* 2010) strongly influence customer brand loyalty. Adding to these views, trust is acknowledged to be part of the critical path towards building brand loyalty (Abreheem and Mohammad, 2012; Lin and Lee, 2012; Kiyani *et al.* 2012; Khraim 2011).

Brands as a self-expression tool – gender differences

It is recognized that brands act as means for expressing self-identity through one's lifetime (Jung and Shung, 2011). Consequently, customers are more likely to purchase products and develop brand loyalty towards brands they perceive as close to their self-image (Rocereto and Mosca, 2012; Moore and Wruster, 2007; Kuenzel and Hilladay, 2010;Archouri and Boulsama, 2010). Thus, the congruence between brand personality and self-concept plays a critical role for building brand loyalty (Park *et al.* 2005). However, Sprott *et al.* (2009) argue that whereas some consumers might develop self-schemas regarding the brands they purchase and like, others might not be likely to develop such schemas. Nonetheless, previous research on the topic emphasizes on the importance of acknowledging gender differences in building brand loyalty (Rocereto and Mosca, 2012). While Godfrey *et al.* (2009) asserts that consumers vary in the manner they perceive companies' relationship-building efforts, Barber (2009) specifies that the

uniqueness of gender groups can act as a direction for their brand judgments, where the self-concept brand-image relationship is gender specific (Lau and Phau, 2009). Further, a study by Hall *et al.* (2000, cited in Barber, 2009) shows that men tend to place higher value on their social and psychological in regard to perceived value of product purchase and/or consumption. In effect, it is critical to gain better comprehension as of the manner in which consumers experience brands (Brakus *et al.* 2009).

Brand Communities

Bergkvist and Larsen (2010) identify a sense of belongingness to a brand community as a critical aspect leading not only to brand loyalty, but to brand love as well. In contrast, Lam *et al.* (2010) presents the concept of social mobility, related to one's intention to change his/her social group, for example, for a higher status group. However, Shen *et al.* (2010) and Lee *et al.* (2011) recognize that participation and satisfaction with virtual brand community, brand communities in general (Heere *et al.* 2011; Labrecque *et al.* 2011; Matzler *et al.* 2011), and personal relationships among community members tend to shape brand loyalty. Additionally, Yeh and Choi (2011) theorize that trust possesses a high degree of influence over one's sense of group belongingness. Moreover, in terms of gender differences, Fan and Miao (2012) assert that women are more likely to base their purchase decisions on member's evaluations and WOM.

Virtual Communities-eWOM and brand loyalty

The advancement of Social Media and Web 2.0 allows customers to build virtual communities, where they exchange information, socialize, and evaluate different products/brands (Christodoulides and Michaelidou, 2010). Thus, Harwood and Garry (2010) introduce the concept of "co-production" between the firm and customers, and the immense opportunity for companies to get better insights into customers' needs and desires (Raihan *et al.* 2013). Although the term Social CRM is relatively new to the literature (Pavicic *et al.* 2011), academics agree on Social Media as being a useful tool for building brand loyalty (Mosadegh and Behboudi, 2011). According   to sociolinguistic theory gender groups demonstrate different communication styles, as women visit virtual communities with the purpose of offering and/or receiving support (Yates, 2001, cited in Awad and Ragowsky, 2008). Moreover, women exert higher levels of e-loyalty with the presence of enhanced social presence (Awad and Ragowsky, 2008; Shen *et al.* 2010). The raise of Social Media has lead to the establishment of Social Media Sites, allowing customers to disseminate brand related information and eWOM (Themba and Mulala, 2013). In particular, academics pay a considerable attention to the effect of eWOM on customer purchase decisions and development of brand loyalty (Chu and Kim, 2011; Chatterjee, 2011; Muntingua *et al.* 2011; Kamtarin, 2012), as Raies *et al.* (2011) state that positive online communities interactions/ties are likely to result in increased brand loyalty.

Research Methodology: Design and Sample

The academic literature agrees on the main contradictory areas regarding positivist and interpretive philosophy. Whereas quantitative methodology bases findings on already existing literature (Bansal *et al.* 2012), qualitative methods are concerned about getting better comprehension about consumers' perceptions, reasons, concerns, attitudes, emotions, preferences (Bratucu and Bratucu, 2012). While the positivist approach emphasizes on well pre-determined variables Charoenruck (2011) and variable interpretation through questions such as "what" and "where" (Prowse and Camfield, 2013), interpretive research builds findings on thorough comprehension of concepts and subjects (Tewksburry, 2009; Saunders *et al.* 2000). Nevertheless, one of the main drawbacks of qualitative methodology stems from the immense amount of information obtained throughout the research process (Bansal and Corely, 2011), challenging the handling of research issues in marketing-related studies (Gummesson, 2005, cited in Kapoulas and Mitic, 2012).

Given the discussion above, it is considered that qualitative methodology is more appropriate for two main reasons. First, as this research will be conducted for a first time in the Balkan area, it is vital to obtain a thorough comprehension of local customers' attitudes, thus fulfill gaps in the academic literature and offer managerial implications. Second, as qualitative research allows for building relationships between the researcher and participants, this increases the chances for more honest responses, a critical factor for the research findings to be obtained.

Online-based research

Provided the fact that the research will encompass the entire Balkan region it is perceived important to consider the supporting role of online-based research. Even though online surveys are heavily criticized (Mouncey, 2012), Barber *et al.* (2013) assert that in contrast to offline surveys, there is an enhanced link between online surveys and external sources, and the quality of data is superior. On the other side, Duffy *et al.* (2005) stress the importance of utilizing a big enough sample so as to guarantee the resulting sample is representative for the entire population.

Sampling method

As sampling remains the most critical basis for qualitative research (Hooper, 2011), it is decided that the research will adopt two sampling methods.
- **Stratified:** divide the sample into three groups, respectively loyal customers, active Social Media users, and less loyal customers. The main reason for separating participants in three groups is to get a better insight into each group attitudes, feelings, and perceptions, thus enhance academic knowledge and provide managerial implications. Additionally, this approach ensures more accurate illustration of the population (Saunders *et al.* 2000). Finally, data will be gathered through focus groups, where constructive insights can be gained from participants' interactions (Sobreperez, 2008). However, the researcher has

to be careful with certain participants exerting control over others' opinions (Kress and Shoffner, 2007), thus leading to some losing their uniqueness, creativity, and individuality (Boateng, 2012).

- **A mixture of snowball and River Sampling responses** (Terhanian and Bremer, 2012) – initially invitations will be send by the researcher to a selected group of people, who are perceived as truthful for further distribution of the survey questions.

Conclusion

Analysis of academic literature demonstrates that nowadays the notion of brand loyalty is a critical factor that distinguishes successful companies. In particular, only when brands go further to emphasize on symbolic (intangible) benefits, they can benefit from loyal customer base (Power and Hauge, 2008). Furthermore, companies need to recognize that customers develop a sense of loyalty towards brands they perceive close to their self-image (Park *et al*. 2005), as the existence of gender differences is acknowledged to play a key role (Rocereto and Mosca, 2012). By understanding the influence of brand communities in the context of Social Media for building brand loyalty (Shen *et al*. 2010; Lee *et al.* 2011), companies can extend their opportunities to unearth new paths towards customers' hearts.

The prior is essential given the fast pace with which Social Media develops to reach a phase where academics and practitioners discuss concepts such as "eWOM" (Themba and Mulala, 2013), "Web 2.0" (Christodoulides and Michaelidou, 2010), "co-production within Social Media" (Harwood and Gary, 2010), and "Social CRM" (Pavicic *et al*. 2013). Therefore, it is considered vital to examine current state of knowledge with respect to building brand loyalty through Social Media within the Balkans in order to fill gaps in academic literature and provide insights to companies operating in this region.

Reference list

Aaker, D. (2012) Win the brand relevance battle and then build competitor barriers. *California Management Review*. 54 (2) p. 43-57

Abreheem, A. and Mohammaad, S. (2012) The effect of brand trust and perceived value in building brand loyalty. *International Research Journal of Finance and Economics.* 86 p. 111-126

Archouri, M. and Boulsama, N. (2010) The effect of the congruence between brand personality and self-image on consumer's satisfaction and loyalty: a conceptual framework. *IBIMA Business Review*. p.1-16

Awad, N. and Ragowsky, A. (2008) Establishing trust in electronic commerce through online word of mouth. *Journal of Management Information Systems*. 24 (4) p. 101-121

Bansal, P. *et al.* (2012) What's different about qualitative research? *Academy of Management Journal*. 55 (3) p. 509-513

Bansal, P. and Corley, K. (2011) The coming of age for qualitative research: embracing the diversity of qualitative methods. Academy of Management Journal. 54 (2) p. 233-237

Barber, N. (2009) Wine consumers information search: gender differences and implications for the hospitality industry. *Tourism and Hospitality Research*. 9 (3) p. 250-269

Barber, T. *et al.* (2013) Moving an established survey online-or not? *International Journal of Market Research*. 55 (2) p. 187-199

Bergkvist, L. and Larsen, T. (2010) Two studies of consequences and actionable antecedents of brand love. *Brand Management*. 17 (7) p. 504-518

Boateng, W. (2012) Evaluating the efficacy of focus group discussion (FGD) in qualitative social research. *International Journal of Business and Social Science*. 3 (7) p. 54-57

Brakus, J. *et al.* (2009) Brand experience: what is it? How is it measured? Does it affect loyalty? *Journal of Marketing*. 73 (3) p. 52-68

Bratucu, G. and Bratucu, T. (2012) Using qualitative research in educational marketing. *Bulletin of the Tansilvania University of Brasov*. 5 (1), p. 19-24

Chahal, H. and Bala, M. (2010) Confirmatory study on brand equity and brand loyalty: A special look at the impact of attitudinal and behavioral loyalty. Cited in: Smith, T. (2012) The personality trait predictors of brand loyalty. *Academy of Business Research Journal*. 3 p. 6-20

Chan, C. *et al.* (2012) Identifiable but not identical: combining social identity and uniqueness motives in choice. *Journal of Consumer Research*. 39 (3) p. 561-573

Charoenruck, D. (2011) Communication research methodologies: qualitative and quantitative methodology. *Qualitative and Quantitative Methodologies*. p. 1- 17

Chatterjee, P. (2011). Drivers of new product recommending and referral behavior on social network sites. *International Journal of Advertising,* 30 (1) p. 77-101

Christodoulides, G. and Michaelidou, N. (2010) Shopping motives as antecedents of e-satisfaction and e-loyalty. *Journal of Marketing Management*. 27 (1-2) p.181-197

Chu, S. and Kim, Y. (2011) Determinants of consumer engagement in electronic word-of-mouth (eWOM) in social networking sites. *International Journal of Advertising*. 30 (1) p.47-75

Duffy, B. *et al.* (2005) Comparing data from online and face- to-face surveys. *International Journal of Market Research.* 47 (6) p. 615–639

Fan, Y. and Miao, F. (2012) Effect of electronic word of mouth on consumer purchase intention: the perspective of gender differences. *International Journal of Electronic Business Management.* 10 (3) p. 175-181

Gecti, F. and Zengin, F. (2013) The relationship between brand trust, brand affect, attitudinal loyalty and behavioral loyalty: a field study towards sport shoes consumers in Turkey. *International Journal of Marketing Studies*. 5 (2) p. 111-119

Godfrey, A. *et al.* (2011) Enough is Enough! The fine line in executing multichannel relational communication. *Journal of Marketing*. 75 p. 94-109

Gummesson, E. (2005) Qualitative research in marketing: road map for wilderness of complexity and unpredictability. Cited in: Kapulas, A. and Mitic, M. (2012) Understanding challenges of qualitative research: rhetorical issues and reality traps. *Qualitative Research Methods: An International Journal* 15 (4). 354-368

Jung, J. and Shung, D. (2011) Brand equity of luxury fashion brands among Chinese and U.S. young female consumers. *Journal of East-West Business.* 17 (1) p. 48-69
Hall, J. *et al.* (2000) Gender differences in a modified perceived value construct for intangible products. Cited in: Barber, N. (2009) Wine consumers information search: gender differences and implications for the hospitality industry. *Tourism and Hospitality Research*. 9 (3) p. 250-269

Harwood, T. and Garry, T. (2010) "It's mine!"-Participation and ownership within virtual co-creation environments. *Journal of Marketing Management*. 26 (3-4) p. 290-301

Heere, B. *et al.* (2011) Brand community development through associated communities: grounding community measurement through Social Identity Theory. *Journal of Marketing Theory and Practice*. 19 (4) p. 407-422

Hooper, C. (2011) Qualitative in context. *Journal of Advertising Research*. 51 p. 163-166
Kamtarin, M. (2012). The effect of Electronic Word of Mouth, Trust and Perceived Value on behavioral Intention from the Perspective of Consumers. *International Journal of Academic Research in Economics and Management Sciences.* 1 (4) p. 56-66.

Khraim, H. (2011) The influence of brand loyalty on cosmetics buying behavior of UAE Female Consumers. *International Journal of Marketing Studies*. 3 (2) p. 123- 133

Kiyani, T. *et al.* (2012) The relationship between brand trust, customer satisfaction, and customer loyalty (evidence from automobile sector of Pakistan). *International Journal of Contemporary Research in Business*. 4 (1) p. 490-502

Kuenzel, S. and Halliday, S. (2010) The chin of effects from reputation and brand personality congruence to brand loyalty: the role of brand identification. *Journal of Targeting, Measurement and Analysis for Marketing*. 18 (3/4) p. 167-176

Labrecque, L. *et al.* (2011) Exploring social motivations for brand loyalty: conformity versus escapism. *Journal of Brand Management*. 18 (7) p. 457-473

Lam, S. *et al.* (2010) Resistance to brand switching when a radically new brand is introduced: a Social Identity Theory perspective. *American Marketing Association*. 74 (6) p. 128-146

Lau, K. and Phau, L. (2009) Impact of gender on perceptual evaluation for prestige brands. *Brand Management*. 17 (5) p. 354-367

Lee, D. *et al.* (2011) The impact of online community type on consumers' community engagement behaviors: consumer-created vs. marketer-created online brand community on on-line social networking websites. *Cyberpsychology, Behaviors, and Social Networking*. 14 (1-2) p. 59-63

Lin, M. and Lee, C. (2012) The influence of website environment on brand loyalty: brand trust and brand affect as mediators. *International Journal of Electronic Business Management*. 10 (4) p. 308-321

Matzler, K. *et al.* (2011) Personality, person-brand fit, and brand community: an investigation of individuals, brands, and brand communities. *Journal of Marketing Management*. 27 (9-10) p. 874-890

Moore, D. and Wruster, D. (2007) Self-brand connections and brand resonance: the role of gender and consumer emotions. *Advances on Consumer Research*. 34 p. 64-66

Mosadegh, M. and Behboudi, M. (2011) Using social network paradigm for developing a conceptual framework for CRM. *Australian Journal of Business and Management Research, 1 (4) p 63-71*

Mouncey, P. (2012) Opening up access to data. *International Journal of Market Research*. 54 (6) p. 727-734

Muntingua, D. *et al.* (2011). Introducing COBRAs; Exploring motivations for brand-related social media use. *International Journal of Advertising.* 30 (1) p. 13-46.

Park, S. *et al.* (2005) Congruence between brand personality and self-image, and the mediating roles of satisfaction and consumer-brand relationship on brand loyalty. *Asia Pacific Advances in Consumer Research.* 6 p.39-45

Pavovicic, J. *et al*. (2011) Customer knowledge management: towards Social CRM. International *Journal of Management Cases*, 13 (3) p. 203-209

Power, D. and Haug,A. (2008) No man's brand-brands, institutions, and fashion. *Growth and Change*. 39 (1) p. 123-143

Prowse, M. and Camfield, L. (2013) Improving the quality of development assistance: what role for qualitative methods in randomized experiences? *Progress in Development Studies.* 13 (1) p. 51-61

Raihan, N. *et al.* (2013) Social Media: an emerging dimension of marketing communication. *Journal of Management and Marketing Research.* 12 p.4-11

Raies, K. *et al.* (2011) Brand loyalty intention among members of a virtual brand community: the dual role of commitment. *Recherche et Applications en Marketing.* 26 (3) p. 23-41

Rocereto, J. and Mosca, J. (2012) Self-concept, gender and product type: an investigation on brand loyalty. *Journal of Business & Economics Research*. 10 (1) p. 25-36

Rossiter, J. and Bellman, S. (2012) Emotional branding pays off: how brands meet share of requirements though branding, companionship, and love. *Journal of Advertising Research.* 52 (3) p. 291-296

Saunders, M et al. (2000) *Research Methods for Business Students.* 2nd edition Pearson Education. Edinburg Gate: Harlow

Shen, Y. *et al.* (2010) Virtual community loyalty: an interpersonal-interactive perspective. *International Journal of Electronic commerce*. 35 (1) p. 49-73

Sprott, D. *et al.* (2009) The importance of a general measure of brand engagement on market behavior: development and validation of a scale. *Journal of Marketing Research.* 46 (1) p. 92-104

Terhanian, G. and Bremer, J. (2012) A smarter way to select respondents for surveys. *International Journal of Market Research.*54 (6) p. 751-780

Tewksbury, R. (2009) Quantitative versus qualitative methods: understanding why qualitative methods are superior from criminology and criminal justice. *Journal of Theoretical and Philosophical Criminology*. 1 (1), p. 38-58

Themba, G. and Mulala, M. (2013) Brand- related eWOM and its effects on purchase decisions: an empirical study of University of Botswana students. *International Journal of Business and Management.*8 (8) p. 31-40

Tiwary, A. and Roy, D. (2012) Measuring brand strength: concept and mobile handset study. *Mobile Marketing Association*.7 (3) p. 38-55

Tsiotsou, R. (2013) The role of brand relationship and tribal behavior on brand loyalty. *American Marketing Association.* 24 p. 366-374

Vivek, S. *et al.* (2012) Customer engagement: exploring customer relationships beyond purchase. *Journal of Marketing Theory and Practice*. 20 (2) p. 127-145

Voss, G. *et al.* (2010) How complementarity and substitution alter the customer satisfaction-repurchase link. *Journal of Marketing*. 74 (6) p. 111-127

Yates, S. (2001) Gender, language, and CMC for education. Cited in: Awad, N. and Ragowsky, A. (2008) Establishing trust in electronic commerce through online word of mouth. *Journal of Management Information Systems*. 24 (4) p. 101-121

Yeh, Y. and Choi, S. (2011) Mini-lovers, maxi mouths: an investigation of antecedents to eWOM intention among brand community members. *Journal of Marketing Communications.* 17 (3) p. 145-162

Yousaf, U. *et al.* (2012) Studying brand loyalty in the cosmetics industry. *Scientific Journal of Logistics.* 8 (4) p. 327-337

E-commerce in Kosovo: Consumer Perception and Adaptation Challenges

Kreshnik Halili[1] and Dr. Anastasios Karamanos[1]

[1] CITY College, An International Faculty of the University of Sheffield
khalili@city.academic.gr

Abstract. The objective of this research is to understand the situation of e-commerce in Kosovo by analysing the consumer perception and adaptation challenges. Existing literature is analysed to construct theory about e-commerce. The theory explores different factors such as security, legislation, trust, awareness, supply-chain and design that influence consumer perception and adaptation challenges. The research is trying to validate theory by using quantitative method. The initial results validate a relatively low influence of factors in consumer perception to e-commerce. The low influence is due to relatively small sample and no previous research in area of e-commerce conducted in Kosovo. Further, the research has practical implications that are related to security of e-commerce, which from existing reports is thought to be critical point for e-commerce diffusion in Kosovo. However, the research finds that consumers in Kosovo are price oriented and that any attractive deal offered by the businesses through e-commerce will overcome the security concern of consumers. Therefore, future researches about e-commerce in Kosovo should consider pricing strategies and their effect on e-commerce consumer perception.

Keywords: *e-commerce consumer behavior, information security, consumers in Kosovo, e-commerce legislation*

1   Introduction

The Internet has changed the way people live. In this information age, de facto people depend on the Internet, i.e. it's becoming a background of our life (Bargh et. al 2004). The physical world we know is transforming into digital. According to Boston Consulting Group report, Internet Economy will reach a total of 4.1 trillion U.S Dollars by 2016. This effect of Internet economy has also changed the consumer behavior especially when purchasing goods or services. The e-commerce has created new opportunities for both consumers and companies. While the e-commerce diffusion is increasing and in some countries is becoming an important factor to economy, e-commerce in Kosovo has low development due to many reasons. In many cases security is one of the reasons why e-commerce in Kosovo has stagnated which could be an assumption. Or, the low development of e-commerce in Kosovo has to do with low innovation from local organizations willing to venture in e-commerce. The purpose of this paper is to understand the actual situation of e-commerce in Kosovo by measuring different areas of e-commerce according to the existing literature. The objectives of

this paper are as follows:

- To analyze the consumer perception to e-commerce in Kosovo and identify key factors
- To understand adaptation challenges of companies or entrepreneurs planning to introduce e-commerce
- What future changes are required to do in Kosovo in order to create better environment for new entrants in e-commerce business
- To create the basis for future researches in e-commerce in Kosovo

Until now, there are no academic papers available that analyses the situation of e-commerce in Kosovo, this paper marks the first academic contribution in this matter. In addition, the paper serves as a guide for future development of e-commerce in Kosovo and may contribute to policy makers as reference.

Will start with analysis of e-commerce definition as such and let us see what does –e-commerce means in practice and move on with adaptation challenges and consumers perception.

Accroding to Henderson et al. (2004) every purchase of services or goods over the Internet falls into e-commerce business. Similarly, Petrovic and Kovacevic (2011) argue that besides the purchase, e-commerce also covers distribution of goods and services over the Internet or electronic communication channels. Contradictory, Belkhamza and Azizi (2009) suggest that e-commerce is a new form of exchange between two parties that never met. These two definitions may be in contrary, but in practice both are valid as in e-commerce services and goods are purchased in the Internet where seller and buyer never met. Ngai and Wat (2002) argue that there is no universal definition of e-commerce, but they follow Kalakota and Whiston (1997) definition of e-commerce that is based in following:

- Communication Perspective: e-commerce is delivery of services or products via telephone lines, computer network or other means
- Business Process Perspective: e-commerce is automation of business transaction through use of technology
- Service Perspective: e-commerce is a tool to cut service costs of companies whilst increasing quality and speed of service delivery
- Online Perspective: e-commerce provides capability of buying and selling products or information in the Internet.

Karakota and Whiston (1997) definition brings different perspectives in terms of e-commerce, however it may not always be in line with latest academic definitions as before. This could be the reason that during the 90's the e-commerce was still seen from different perspective. Though, the only comparable perspective of e-commerce from Karakota and Whiston could be considered the Online Perspective that addresses the buying and selling over the Internet. On the other hand, The Organization for

Economic Co-operation and Development (2000) argues that there is no widely accepted definition of e-commerce as it may be seen from different perspectives.

2   Literature Review

To understand two parts of the "coin" we have analyzed literature from two perceptives (a) to understand the adaptation challenges by the companies and (b) consumer perception to e-commerce. The literature review will help to create theory about e-commerce.

2.1 E-commerce as part of innovation cycle

The adaptation of e-commerce has to do with vision of the managers leading the organization and their entrepreneur spirit of how far could see the future in terms of technology development and innovation (Molla et al, 2006). In fact adaptation of e-commerce is an example of innovation (Kendall et al., 2001). In addition Molla stated, (2006) "Innovation theory suggests that the adoption process in organizations passes through several phases, which are typically labeled initiation, decision, and implementation". These phases are important to understand the adaptation or starting a new e-commerce business and entrepreneur mindset. During the initiation the managers start to push the idea of e-commerce. Initiation phase is characterized with a level of uncertainty, which is a typical situation in corporate venturing as well (Daft and Lengel, 1986). The next phase is decision in which the entrepreneur is convinced to embrace e-commerce and have understood the benefits. On the other hand, the decision phase according to Conner "Change Transition Model" can be part of uninformed optimism a situation where entrepreneur make decision without having complete information (Rusell, 2006).
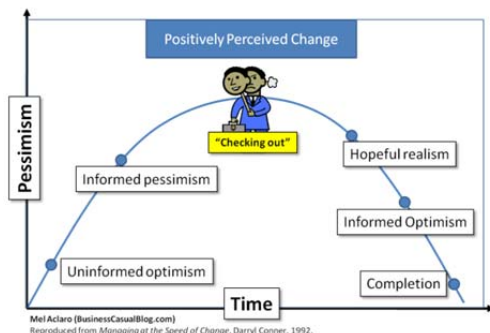
Fig. 1 Conner Change Model (Source:
http://www.businesscasualblog.com/2008/10/informed-pessimism-always-follows-uninformed-optimism.html)

This situation can be very delicate as entrepreneurs may follow a wrong path that may lead them to checkout, a phase of model in which entrepreneur cease to follow the strategy. The last phase implementation is more or less straight forward as entrepreneur has understood the risk and according to Conner model has now moved to

informed optimism a phase where entrepreneur is aware of change or in our case adaptation of e-commerce. It's worth using Conner's model to understand entrepreneurs mind when considering e-commerce especially in Kosovo where many of them are "traditional entrepreneurs" used of exploiting traditional business ideas such as trading goods or offering basic services e.g. cleaning companies.

While the process may seem simple, adaptation of e-commerce is challenging especially if a business is solely dependent on e-commerce. This proved also during 2002 e-commerce boom that ended with many bankruptcies (Barsauskas, Sarapovas, Cvilikas, 2007). Therefore, a careful planning with sufficient resources has to be considered before entering e-commerce business.

In addition the model above, Molla et al. (2006) has created a conceptual framework for e-commerce adaption based on 3 stages followed by key influencing factors such as Contextual, Organizational, Managerial and E-commerce specific. (Fig 2.)

According to the model, there are 3 stages of e-commerce that also create an evolution process. Stage 1: The entrepreneur or company has started with a simple website just to have an Internet presence, but based on the benefits perceived through influencing factors has moved to the Stage 2: The website has some interaction features where visitors can search products and get more details. In Stage 3: Transaction processing is integrated in the e-commerce that gives visitors online payment capabilities. In all three stages, influencing factors are key to move the evolution of e-commerce.
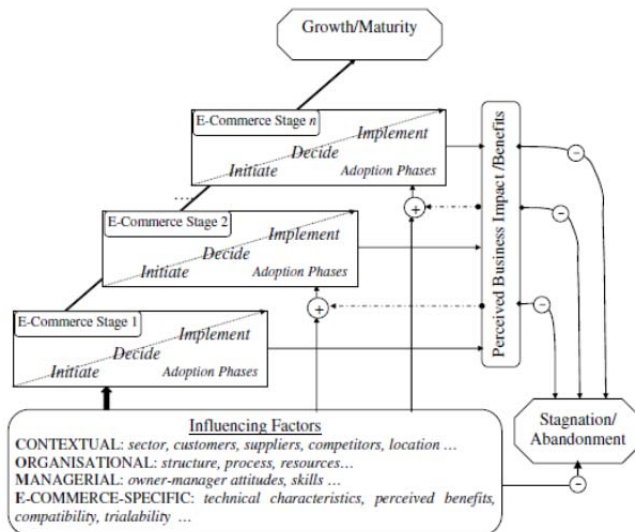
Fig. 2 A conceptual framework for the e-commerce adoption process (Molla et al, 2006)

Two models Conner and Molla are of two different natures, but in combination give us a great understanding on what challenges entrepreneurs have to go through before entering or adopting e-commerce.

2.2   E-commerce and development of Information Technology

Many studies have related the development of Information Technology (IT) with performance of company. For instance, (Hitt and Brynjolfsson, 1996, Tallon et, al 2000) have analyzed that investment in IT will increase productivity of company, but not necessary the profitability. On the other hand, some studies suggest that IT effect in financial performance is higher if the strategy of the company is aligned closer with IT (Li and Ye, 1999). However, Lee (2001) has created a model that proves that well managed IT can bring many benefits to the company as well as profitability. She adds that IT can have a great impact in quality of processes of company and reflection in efficiency.

   According to Koufaris (2002) companies that are heavily utilizing IT are ready to easily adopt e-commerce. He adds that internal IT users are similar to consumers. Pavlou and Fygonson (2006) state that e-commerce encompasses IT elements such as technology and software with marketing elements. Gefen and Straub (2000) take different approach by linking the IT and e-commerce according to consumer behavior. Processes facilitating purchases are more complex and require more planning.

2.3   Security of E-Commerce

Businesses willing to adapt e-commerce should have comprehensive security of systems (Al-Slamy, 2008). Securing e-commerce channel proved to be very challenging for companies. The investment required to achieve a certain level of assurance of e-commerce are in many cases high. The security of e-commerce relies on features in different levels of the system (Al-maaitah, 2011). Maintaining security of e-commerce will increase the trust of e-consumers.

   According to Blythe (2008) there are four levels of e-commerce security. These four levels can be considered to secure e-commerce in order to increase consumer confidence. The security as suggested by Blythe (2008) has to do with online contracts. The following are the levels to that could be incorporated as part of online contracts to secure e-commerce
   1. The very basic one is the "I Agree" button that ensures that consumers understand and agrees to the terms of online contract
   2. In step consumers is required to enter a secret in form of password or credit-card in order to confirm the purchase
   3. Biometrics are yet another form of security that could be used to secure e-commerce.
   4. Digital Signature as one of the most advanced security measures to secure e-commerce. However, incorporating digital signatures requires a well-established process within company.

   Antoniou and Batten (2011) argue that before securing e-commerce the first step is to identify the potential risks and then treat them accordingly. Nevertheless, like Blythe (2008) they share the opinion that digital signatures remain as the best option to secure e-commerce.

2.4   E-Commerce Legislation Framework Challenges

Apart from the vision of the company or security of e-commerce, the other aspect towards adoption is the legislation. According to UN Trade and Investment Division (2007), e-commerce main character is geographic independence. At the same time this characteristic makes it difficult to determine jurisdiction.

Many countries have struggled to develop a legislation framework that addresses the business of e-commerce. Lack of a mature legislation framework has created another ambiguous situation for companies trying to adopt e-commerce. Oxley and Yeung (2001) argue that e-commerce legislation is very important to create a stable e-commerce environment. The e-commerce legislation is very important to increase consumer confidence as well. The success of e-commerce depends on the government presence to offer the necessary legislative instruments in order to increase credibility of online markets (Oxley and Yeung, 2001).

U.S and EU took the first steps to regulate e-commerce legislation in their territories. U.S as the pioneer in e-commerce took the first step in 1997 by introducing "Framework for Global Electronic Commerce,". U.S wanted private sector to be more active in order to regulate e-commerce. Farrell (2003) states that while U.S wanted to quickly organize e-commerce, EU was more cautious about the policies. EU expressed many concerns in privacy of data flowing internationally. However, in 1997 EU adopted "European Initiative on Electronic Commerce" in order to create a better market for e-commerce. The regulation had the following objectives:

1.   To provide widespread, affordable access to the infrastructure, products, and services needed for electronic commerce;
2.   To create a favorable regulatory framework for electronic commerce;
3.   To foster a favorable business environment for electronic commerce; and
4.   To ensure that the global regulatory framework for electronic commerce is coherent and compatible with that of the European Union.

Pearson (2000) argues about a challenge in EU e-commerce legislation initiative due to diverse local legislation of the members. Pearson (2000) adds that in such environment consumers have many legitimate concerns when using e-commerce. However, EU regulation about e-commerce still has its own benefits because it harmonizes regulation for Digital Signature and Encryption. Further, in year 2000, EU released another directive 2000/31/EC that specifically deals with e-commerce. As stated on the point 8

"The objective of this Directive is to create a legal framework to ensure the free movement of information society services between Member States and not to harmonise the field of criminal law as such".

According to Trivun and Siljacic (2007), EU directive does not focus in e-business, but rather on information society service providers which is considered the most important part of the directive. The directive requires member states to establish and legally acknowledge contracts established by electronic means. Electronically con-

cluded contracts by information society service providers are obliged to inform customers with the following information (Trivun and Siljacic, 2007)

- Different technical steps to follow to conclude the contract;
- Whether or not concluded contract will be filed by the service provider and whether it will be accessible;
- The technical means for identifying and correcting input errors prior to the placing of the order;
- The languages offered for the conclusion of the contract, and codes of conduct the service provider subscribes to, and information on how those codes can be consulted electronically.

The development of this directive has created a great foundation for countries in EU willing to build e-commerce in their territory. However, the directive as such is not limited just to EU members, but other countries that are in the process of becoming members could adopt the directive. For example many countries South Eastern Europe have adopted directive 2000/31/EC in the corresponding national e-commerce legal framework. Examples of these countries are Serbia, Macedonia, Albania, and Bosnia.

2.5 Digital Signatures Law

Previously we gave an introduction to digital signatures and their technical ability to ensure integrity and authenticity of the secure electronic communication. It was also mentioned that digital signatures are equal to written signatures though this has to be supported legally. In a study performed by Kuno and Barchero (2000) it can be suggested that each country implements digital signatures laws differently. These laws tend to have different approaches towards digital signatures. The researchers of the study have stated 3 approaches as following:

1. "The first, a minimalist approach, aims to facilitate the use of electronic signatures generally, rather than advocate a specific protocol or technology. "
2. "The second approach tends to be more prescriptive. Here the motivation often stems from a desire to establish a legal framework for the operation of PKIs - whether or not other forms of secure authentication are included or permitted - as well as a reflection of form and handwriting requirements that apply in the offline world."
3. "Some jurisdictions have also begun to realize that the first two approaches are not necessarily mutually exclusive, and so have adopted a third approach. The result has been a "two-tier" approach representing a convergence and synthesis of the two approaches."

2.6   E-commerce and Supply Chain

As a new form of business, e-commerce had its impact also in supply chain. While supply chain is a very wide discipline and requires more time to analyze it, we have put our focus in general terms to understand the difference between the conventional and e-commerce supply chain specifically in B2C market.

The supply chain of e-commerce is different from traditional purchasing in many forms. Mentzer et. al. (2001) defines supply chain as connection of two or more organizations with a purpose to deliver products or services to the end customers. Other researcher have given supply chain a broader definition which is more layered in terms of different processes such design of product, procuring the raw materials and turn them into final products to be delivered to the end-customer (Lee and Billington, 1993). Fig 4 depicts a typical supply chain in which process starts from suppliers up to retailers. It's worth noting; that in this scenario there two flows to be considered the forward flow that presents the actual goods and reverse flow that represents information. For the purpose of this paper we will refer to Mentzer et. al (2001) definition. Hsieh (2012) argues that e-commerce has created different opportunities suppliers.
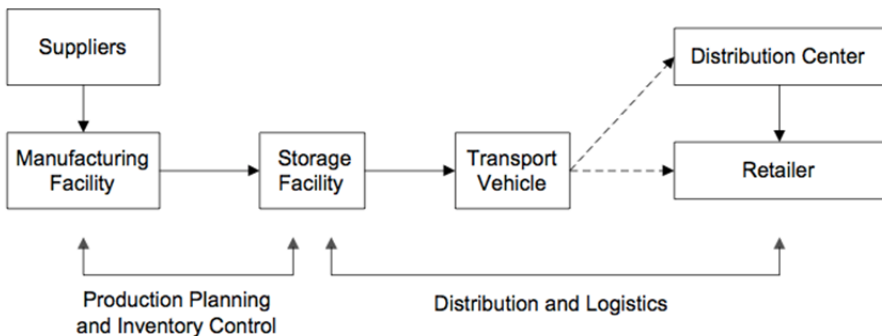
Fig. 4 Typical Supply Chain

In relation to e-commerce and supply chain, most of articles emphasize the aspects of the distribution and logistics that deals with movement of goods.

In a conventional distribution and logistics, goods are distributed to retail shops where customers purchase them. On the other hand, in case of e-commerce, customers purchase online and wait for the goods to be delivered at their address. Hsieh (2012) explains both situation form the real life example by analyzing a women shoe shop. Women shop keeps a high number of varieties of shoes in stores. Shoes take lot of space which is expensive especially in premium retail shops. This situation leads retailers to 1) keep a limited number of shoes in stock and 2) to organize daily delivery to keep the stock. This results in higher costs to manage the stock and in some cases unable to serve the consumers.

The same situation is analyzed by Hsieh (2012) in e-commerce or online retailers where stock is organized differently. For instance an online retailer creates a central warehouse where are goods are distributed to consumers. In this situation, online

retailer can keep higher amount of goods in warehouse because the cost to store them is lower than a high street retailer. Rahman et. Al (2001) argues that while online retailers have the cost advantage for central warehouse, they may lack consumer experience of retail shop where products can be analyzed before purchase. A comparison of two situations is depicted in Fig 5.

Likewise, in case consumer is not happy with product purchased online, then they have to send it back to which causes additional cost and consumer dissatisfaction. According to Lewis (2001) overcoming these situation and create better experience for consumers, online retailers must have close interconnection with logistics companies in order to provide consumer with more information about the delivery of product. Bringing these challenges in Kosovo is even more complicated. At this point, Kosovo is still struggling to have a harmonized address system that until now is still in progress.
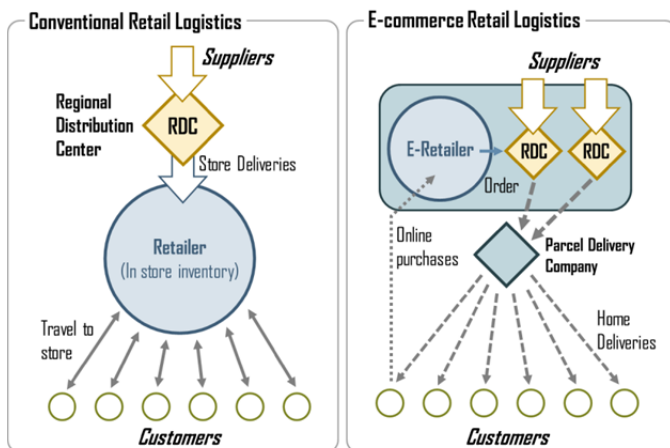
Fig. 5 Comparison of Conventional and E-commerce retail Logistics (Source: http://people.hofstra.edu/geotrans/eng/ch5en/conc5en/ecommercelog.html)

3   Consumer Perception

Implementation of e-commerce in Kosovo is challenging compare to the rest of the region, especially if compared to Western Europe. But what makes Kosovo different in e-commerce is consumer perception to it. Currently the only active e-commerce service in Kosovo is e-banking that has been introduced by many banks since 2006. Today all commercial banks in Kosovo offer e-banking services (The Kosovo Banker, 2012). The adoption starter slow, but nowadays is a banking service that many cannot live without. According to U.S Commercial Agency, Post Telecom of Kosovo and Kosovo Energy Corporate offer payment of bills online.

However, there is slow implementation or introduction of retail e-commerce compare to the rest of European countries. According to "Forrester Research" e-tail sales in Europe during 2012 reached 103.1 billion Euros an annual growth of 10% a clear sign that e-commerce in Europe is flourishing and consumer loyalty is there. In Koso-

vo situation is different and in contradiction to the trends of Europe. In a survey done by The Kosovo Association of Information and Communication Technology (STIKK) consumers are not ready yet to purchase online (Fig. 5). 40% of responders will not use their credit cards online while only 27% will use it.
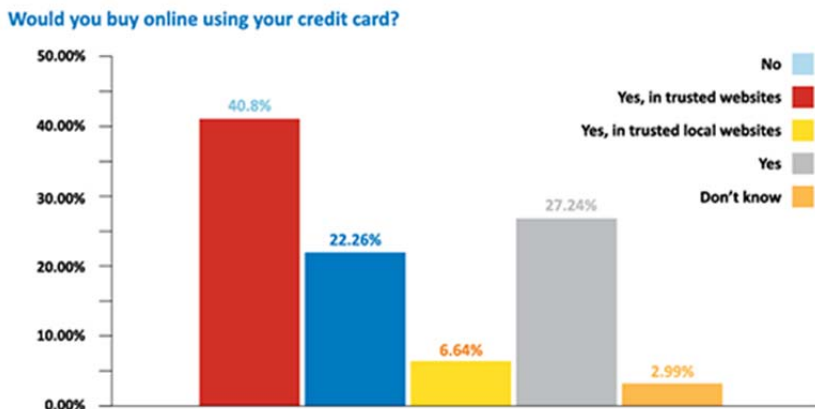
Fig. 6 Online Shopping Confidence (Source: STIKK)

According to report, many responders have stressed security as the main reason for not using credit card online. Threats to e-commerce sites are increasing and the risk of monetary loss in online purchase has driven consumers away from adaptation of e-commerce (Hansen, 2001) hence the right security controls have to be implemented.

On the other hand, Horbs, Kelly and Hofman (2005) argue that retail e-commerce must have proper recovery procedures in place in case the consumers are affected by loss or bad purchase experience. The recovery procedures must be of any forms including refund up to verbal apologies. This could be a simple form to gain consumer trust in case of loss or bad experience in e-commerce.

Persefoni, Stavros, Giannoula and Anastasio (2009) for their study have concluded that besides other factors such as security as mentioned above "the main factors influencing consumer behavior [perception] in e- shopping are culture, trust in transactions and vendors, usefulness and use of e-banking". They came to this conclusion after analyzing consumer behavior of women in Baltic and Balkan countries. Kosovo as part of Balkan can best be compared to this study especially consumer perception.

Region - Do you think that e-commerce is useful for you?		
	Baltic	**Balkan**
Yes	92.05	60.22
No	7.95	39.78

Fig. 7 Usefulness of e-commerce (Source: Scientific Bulletin – Economic Sciences, Vol. 10 (16)/ Issue 2)

In question "Do you think that e-commerce is useful for you?" (Fig. 3) 60% of women in Balkan answered "Yes" compare to Baltic that scored 92%. This may sound optimistic at first instance, but 39% of women in Balkan answered "No" to the same question compared to Baltic region that answered only 7%. This paper correlates also to STIKK survey in Kosovo as part of Balkan where consumers still do not use e-commerce for various reasons.

There are many aspects to be analyzed in order to understand consumer perception towards e-commerce. Considering the nature of e-commerce where consumers cannot touch or feel products trust becomes very crucial. Further in paper we will cover the elements that may influence trust.

One theory suggested by Fishbein and Ajzen (1975) also known as "Theory of Reasoned Action (TRA)" suggest a relationship between attitude, behavior and intentions. The TRA further suggest that people make decision based on the best available information to them. The theory further explains the relation between attitude and behavior and factors influenced each other however, quality of information is one aspect that relates to consumer perception to e-commerce. Therefore is important that e-commerce business to provide the necessary information to consumers in order to increase trust. Eid (2011) suggest that e-commerce websites should offer a form of interactive communication with consumers to offer them the necessary information and clarify any perception they may have about e-commerce. Apart the quality of the information the following will cover other aspects that influence consumers trust.

3.1   Consumer Culture in E-commerce

As mention above by Persefoni, Stavros, Giannoula and Anastasio (2009) culture is one of the factors that have a relation to consumer perception to e-commerce.

In general terms, Erez and Early (1993) define culture as a value shared among society or a group of people. On the hand, a definition that is more specific to consumer culture, is defined by McCort and Malhotrs (1993, p 97) as follows:

"Consumer culture is the shared, consumption relevant knowledge system necessary to operate in a manner acceptable to one's society. This knowledge system, through the formation of culturally learned rules of perception and interpretation, imbues objects and behavior with meaning for its members." Further, McCort and Malhorts (1993) point-out that culture has two forms of influences. First, like Erez and Early (1993) in group level culture influences values, norms and symbols. Second, it influences cognitive functions in individual level (Liao et. al, 2008). These are important elements of cultures as they define consumer behavior in e-commerce. Although in the world of globalization it is believed that consumer cultural values have become more homogenized it proved not to be the case (De Mooij, 2004). According, to De Mooij the consumer culture remain to be intact and diverse among different countries. Aldeen, Senkamp and Batra (1999) have created two marketing strategies to address consumer culture, Global Consumer Culture Positioning (GCCP) and Local Consumer Culture Positioning (LCCP). While they study is not directly related to consumer culture behavior, it still emphasizes the importance culture from advertising perspective. According to Aldeen, Steenkamp and Batra (1999) most of

the companies embrace LCCP marketing strategy, as it tends to address specific local cultural aspects of consumers.

The consumer culture was also investigate in a in another study performed by Lightner et. al (2002). The researchers surveyed students from Turkey and U.S and compared the analysis. As well in this study is can been that Turkish students have some concerns of security and trust in e-commerce compare to U.S students. What is notable in this study is the motivation factors to use e-commerce by Turkey students. Special prices that emulate typical Turkey bazars increased the number of e-commerce users as the experience can be related to the real-life purchases. On the other side, U.S students appeared to have higher confidence in e-commerce which according to Lighter et. al (2002) has to do with experience in computer usage.

3.2   Payment Methods in E-commerce

Payment methods in e-commerce could be considered more as technical feature, however, it has a relation with consumer culture as well. For instance consumers in U.S use their credit cards for online payment compare to EU that use debit cards. On the other hand, Asians use check mail or preferably cash payments (Liao et. al, 2008). Many studies (Kshetri 2001, Chau et al. 2002, Bin et al. 2003, Efendioglu and Yip 2004) have raised the issue of payment methods as a factor for declining the diffusion of e-commerce. The payment methods issue is very important factor in development of e-commerce in Kosovo. According to EU Report on Anti-Money-Laundry, Kosovo has a cash-based economy that is still struggling to move to electronic payments. Moreover, Central Bank of Kosovo is trying to introduce new technological means to support electronic payments. In addition, currently in Kosovo there is only one bank (ProCredit Bank) that can process Visa credit/debit payments over the Internet. This paper investigates consumer perception in Kosovo to online payments in contrast to cash payments.

3.3   Cultural Impact – Design, Usability and Format of E-commerce

Liao et. al (2008) suggests that design of e-commerce website have to address the culture of the targeted audience in the particular country. The website that addresses the cultural aspect of audience will increase the usability of the e-commerce thus making it easier for the consumers to purchase goods online. The usability aspects have been analyzed also by Nielsen (2005) study. He finds that Europeans rated with low marks usability of 20 U.S websites. Nielsen (2005) reports, U.S users have difficulty to navigate e-banking websites in Europe. Interesting fact in this study is that all websites had excellent translation. In another study performed by Lo and Gong (2005) it was concluded that design of e-commerce have a link with consumer culture. The researchers analyzed websites in U.S and Chine to determine design elements and perception by consumers. Lo and Gong (2005) argue that e-commerce design has an impact also in increase of consumer trust. For instance in China, e-commerce consumers have indicated that use of active web elements such as animations increase the credibility of the website. However, the researchers suggest their study is limited

because culture is very complex and is difficult to analyze all dimensions. Barber and Brade (1998) argue that success of e-commerce can be achieved only if the interface is designed to meet cultural aspect of target audience. This is clear indication that different cultures have different expectation or behaviors towards the same e-commerce website. Therefore, is significantly important that web developers should take into consideration the cultural aspects when designing e-commerce websites.

4   Research Method Architecture

From literature review we have identified many factors that are related to e-commerce consumer perception and adaptation challenges. This will lead us to generate constructs (Table 1) in order to understand the situation in Kosovo. Further, we create a model consisted of constructs/factors that will act as independent variables and measure their influence on dependent variables. To validate the model we have utilized Structured Equation Model methodology (SEM). For the purpose of this research, SEM methodology will have confirmatory and exploratory nature i.e. to confirm the below model and explore variables for other potential factors. The results will be graphically analyzed in order to better present the distribution and relation of variables. In our case quantification of survey is analyzed with use of statistical tool Microsoft Excel for transformation of data, IBM SPSS for descriptive, Exploratory Factor Analysis (EFA), and IBM SPSS AMOS for Confirmatory Factor Analysis (CFA) and SEM. Besides the measure of the constructs, the purpose of quantification in our research is also to understand different aspects of survey responders including gender, age, profession etc. Likewise, to understand the correlation between variables that may drive us to different dimension of results of the survey.
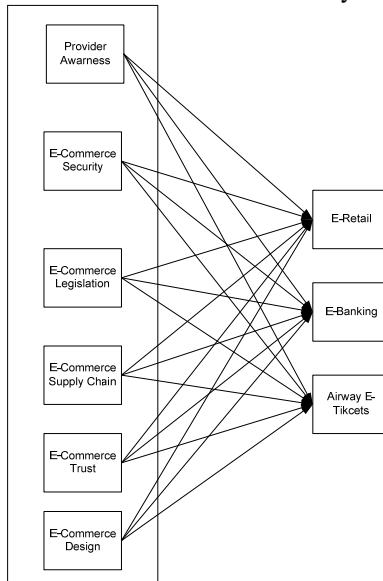
Fig. 8 Research Model

The data collection method utilized in this research is performed with use of a survey questionnaire consisted of three parts. A questionnaire of 36 questions is distributed to the sample as explained in the next section. First part is consisted of question specially tailored to measure constructs in Table 1. Second part is consisted of general questions related to usage of e-commerce by consumers in Kosovo. And in third part are questions related to demographics of responders such as age, gender etc.

Constructs	Definitions	References
E-Commerce Investment	Measure readiness of Kosovo companies to investment in e-commerce	Molla et al. (2006), Lee (2001)
E-Commerce Security	Kosovo consumers perception to e-commerce security	(Blythe, 2008), (Antoniou and Batten, 2011)
E-Commerce Legislations	Kosovo consumers perception to importance of e-commerce legislation	(Kuno and Barchero, 2000), (Pointcheval and Stern ,2000) , Oxley and Yeung (2001)
E-Commerce Supply Chain	Kosovo consumers and companies perception to e-commerce supply chain	(Hsieh, 2012), Lewis (2001)
E-Commerce Customer Trust	Kosovo consumer trust with e-commerce providers	(Polychronidou et. al., 2009), Liao et. al. (2008), (Kshetri 2001)
E-Commerce Design	User Interface of E-commerce website has an impact on consumers trust.	(Polychronidou et. al., 2009), Liao et. al. (2008), (Kshetri 2001)

Table 1 Potential constructs for the research

The survey questions response we have used Likert-type scale format from 1= Strongly Agree to 5= Strongly Disagree. Other general and demographic questions are a mix of different types. However, the focus will mainly be on construct measure questions. The full list of questions and their structure can be found in Annex A of this document.

4.1   Sample

For the purpose of this paper non-probability convenience sampling technique is utilized.

To follow the basic criteria mentioned by Ferber (1977) we created a representative sample as following:

Relevance: The questionnaire is distributed to persons that are relevant to the research who are currently employed and have basic know-how in computers.

Adequate Size: Questionnaire is distributed to 150 employees of Raiffeisen Bank Kosovo, published to Facebook reaching 200 connections and to 30 employees at Innovation Centre Kosovo.

Representative: 123 responses are received in the questionnaire posted in Google Forms i.e. 123/380 = 32% response rate. The responses were anonymous and on voluntary basis hence no consent was requested.

5   Results

As part of the result analysis we have analyzed the demographics of the responders that are very diverse yet targeted for our research. Out of the 123 responders 60.2% are Males and 39.8% Females. In general this is a good gender balance taken from the responses and adequate for further analysis of the results.

The other important aspect for the consumer perception is also the education level of responders. 5.7% are with Secondary school, 49.6% are with Bachelor degree, 43.9% are with Master degree and 0.8% PhD. Here again, we have an indication that our responders are highly educated people. Moreover, 100% of responders have Internet connections while 97.2% hold a credit or debit card. In terms of age, out of 123 responders 82.9% fall between the age of 25-40, 7.3% percent are above 40 and 9.8% are under 25. This is also in-line with average age of population in Kosovo that according to statistical office is 25 years (Bergman, 2010)

From the e-commerce industry point of view we have analyzed the frequencies of four questions in which we have asked the responders about the visit of different e-commerce websites. It can be seen that E-banking is most used e-commerce industry in Kosovo with 85.4% of responders answering with Yes. This also confirms the assumptions as found in the magazine of Bankers Association of Kosovo. The following after e-banking is Online Airplane Tickets purchase with 69.1% responding with Yes. The high number of online airplane tickets is related also with the fact of high number of Kosovars living abroad. Two "cheap flights" airways companies, EasyJet and Wizz Air offer plane tickets at very low price that are very attractive for consumers in Kosovo. In this regards, it could be as well believed that price is a good motivator for consumer to use e-commerce websites. Third e-commerce industry is online purchases or e-tail with 65% answering Yes. This is an interesting fact because currently in Kosovo there is only one online shop. However, speaking to some of responders, they claim that use of international e-tailers specifically www.asos.com is very high though, this has to be further investigated. The last e-commerce industry is online auctions with only 35% answering Yes. The assumption of low responses is that there is no online auction e-commerce website in Kosovo. Below is the table of all response. The industry questions are used as dependent variables in the section, Data Analysis.

		Frequency	**Percentage**
E-Banking Service	Yes/Po	105	85.4
	No/Jo	18	14.6
Online Purchase (Books, Electronics etc.)	Yes/Po	80	65
	No/Jo	43	35

Online Airplane Tickets	Yes/Po	85	69.1
	No/Jo	38	30.9

Table 2 E-Commerce usage by Industry

5.1   Data Analysis

For the purpose of the data analysis we have applied Structured Equation Modeling (SEM) method. The method is a very advanced and requires careful analysis of data. The analyses have are performed with use of IBM SPSS and AMOS software. First to analyze the priori model as suggested in previous section we have used IBM SPSS software to extract the factors from the variables in the survey. From this extraction we applied the Explanatory Factor Modeling (EFA). In below table (Table 3) we can see that not all items have been placed or loaded sufficiently to confirm the proposed model. Therefore, we have had to re-specify the priori model based on the exploratory findings. According, to Heyton et. Al (2004) when using EFA one of the critical part of the process is the decision of retaining factors. In the below output from the SPSS we can see the result is not very well structured and therefore we have to make certain decision before deciding which factors to retain. It is confirmed that the proposed model in the previous section requires some changes. While we have proposed 6 factors to be evaluated the SPSS gave us 8. Therefore, we have carefully analyzed the possible factors to be retained in order to test our model.

When deciding which factors to retain, there are no unique answers or criteria. Carraher and Buckley (1991) argue that current available factor retention criteria produce different results in each scenario. It's also important to researcher to decide if the retained factors are important, make sense and relate to the research question. Heyton et. Al (2004) suggests that keeping too few factors may have a severe impact in the research therefore the researcher has to take this into consideration when deciding about the retained factors. In our model we have used two criterions for factor retention a) by eliminating the items that were loading in two factors and b) eliminating items that have a low level of loading. The outcome of simplified structure produced four valid factors.

	Factor			
	1	2	3	4
Cronbach's Alpha	**809**	**616**	**559**	**660**
Q5	.887			
Q8	.832			
Q10	.666			
Q9	.637			
Q12	.441			
Q19R		986		
Q20		.587		
Q11			.682	
Q6			.541	

Q7R			.494	
Q26				.730
Q18				.696

Table 3 Exploratory Factor Analysis – SPSS Output

In addition, to make sure that retained factors are reliable we have performed another Cronbach's Alpha (CA) test. The purpose of CA test is to ensure the internal consistency of the items that measure a factor or construct (Tavakol and Dennick, 2011). In simple language, to make sure that items within the construct are measuring the same concept. CA is expressed as a number in a scale from 0 to 1. Low scale of CA indicates that items in the construct (factor) are not consistent and therefore could be criticized as not representative and be questioned. Further, there are many discussions among researchers to define the right cut-off level of CA test. Many researchers argue that having a reliable factor, the items should reach a CA of 0.7. However, on the other hand Cortina (1993) argues that CA test should be interpreted with cautions, as there are other elements that may have an impact on the level. In her findings, explains that CA is also affected by the number of items within the factor. Therefore, besides relaying solely in CA number, researcher should investigate the relevancies of the items.

Taking this into consideration, we performed CA reliability test for our retained factors as in Table 3. The test resulted in three factors having a CA higher than 0.6 indicating good reliability of the items and one factor had a CA lower than 0.6 .Factor with lower CA is retained due to the relevance and number of items. Moreover, the factor is measuring the E-Commerce Security, which is very important for the goal of this research.

For next validity measurement, we have utilized Confirmatory Factor Analysis (CFA) to test factor structure of the model that was extracted during the process of Explanatory Factor Analysis (EFA). According to Brown (2006) CFA is part of SEM that addresses measurement part of the model. Measurement is performed by analyzing the observed variables and their relation to latent variables (factors) with calculation of fit-indices. Once the fit-indices reach a cut-off level, then we validate the model fit.

There are many indices to be measured in CFA, however for the purpose of this research we have used goodness-of-fit (GFI) proposed by Joreskog and Sorbom (1993) which is an alternative to Chi-Square that calculates the proportion of variance, root mean square error approximation (RMSEA), proposed by Steiger and Lind (1980) which indicates how well the model fits to population covariance matrix and Comparative-Fit-Index (CFI) proposed by Bentler (1990) which compares X2 of model and null model.

One important note by Brown (2006) is that when applying CFA to measure the model, the research should have a strong theoretical background. In addition, Reise et. al (1993, p554) stated " no CFA model should be accepted or rejected on statistical grounds alone; theory, judgment and persuasive argument should play a key role in defending the adequacy of any estimated CFA model." That said, our proposed CFA model is supported by the existing theory on E-Commerce.

By using these above three fit indices we measure CFA model specified out of the items (observed variables) and factors (latent variables) that are retained during the procedure of EFA. To create the model we have used the SPSS AMOS 20 software package. The first model crated in AMOS is seen in the Fig 12 below. After creating the model and connecting all observed and latent variables we run the estimation to generate the above fit indices. CFA model has four valid factors
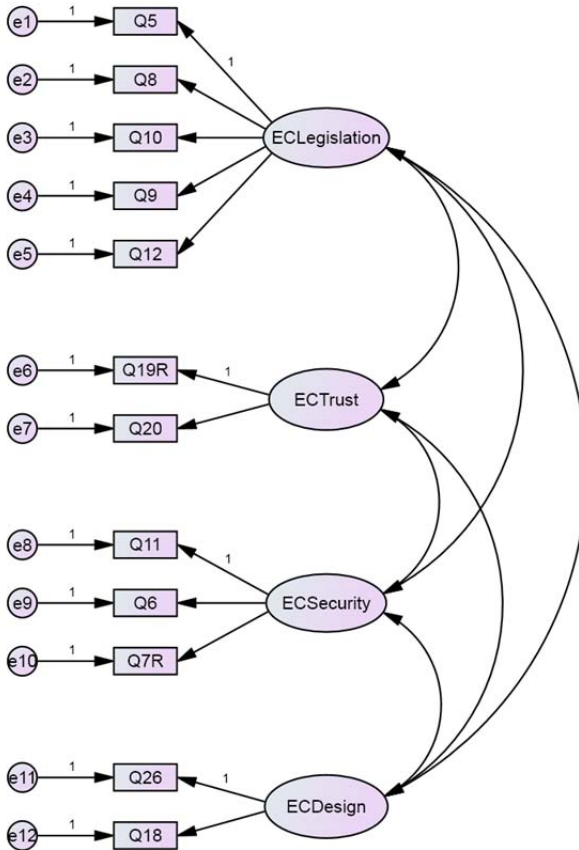
Fig. 9 Confirmatory Factor Analysis Model generated with AMOS

The four factors (ECLegislation, ECTrust, ECSecurity and ECDesign) fit to the purpose of the research. The above model has reached all cut-off level of indices with GFI=.909, CFI=.931 and RMSEA=.072 validating a fit model. However, during the estimation of model in AMOS, it was reported that one of error terms "e12" has a negative variance also known as Heywood case. According to Kolenikov and Bollen (2007) there are many causes for error term negative variances in a model such as outliers, non-convergence, empirical under-identification, structural misspecification

and sample fluctuation. In case of our research we can conclude that we have a sample fluctuation due to the small size N=123.

There are some methods to remediate the problem of negative variances and the researcher has to analyze the possible solutions. One of the solutions is to remove the item or factor that has a negative variance. Bentler and Chou (1987) argue against of removing the item or factor, instead it has to be analyzed and understand the root-cause. The other option is to understand the level of negative variance Being (2007) explains identification of sampling error. However, for our model the best solution to have an admissible model is do delete factor ECDesign. The decision can be taken in basis that even the remaining factors are still important for the E-Commerce consumer perception, if not the most important. Likewise, with removal of ECDesign we have managed to keep the cut-offs of GFI, CFI and RMSEA indices. In addition, have a simple CFA model is very important for specification of structure model.

The next and the final part of the SEM, is specification and validation of the structure model. In our case we have used three exogenous variables as the output from the CFA model, three endogenous variables which are acting as dependent variable and three control variables Gender, Age and Education. With structural model we analyze the relationship between the factors and the effect on usage of e-commerce industries E-tail, E-banking and Airline. Moreover, we identify if factors have any effect on consumers when using e-commerce. The model is generated in AMOS and it looks as in Fig 13.
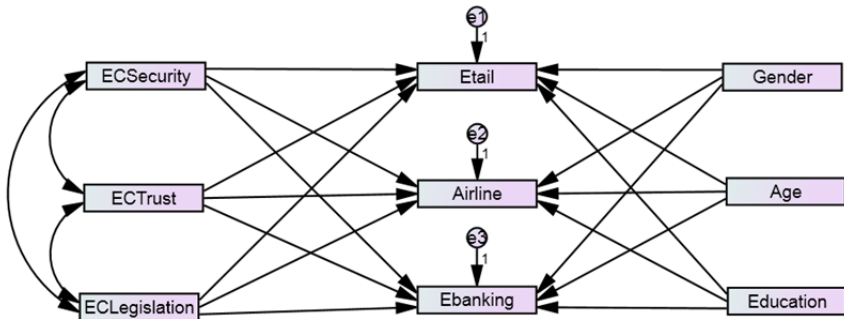
Fig. 10 Structured Equation Model - exogenous and endogenous variables

There are many aspects to understand the model and if it should be accepted or rejected. The absolute fit indices report a Chi-Square= 48.065 with df=15 and p=.000 indicating model is completely adequate. Chi-Square is very sensitive to the sample size therefore we focus in other fit indices to validate the model.   Starting from the fit indices, the model has reached moderately cut-off level of previously discussed fit indices with CFI=.801, GFI=.920 and RMSEA=.134. Taking this into consideration we can conclude that the model is fit with minor deviation. However, we want to evaluate other aspects of the model that may have an implication on our research.

Regression analyses are performed with priori model and assuming researcher has already performed some investigation and tries to understand the effect of variables

on each other. This is also important for our research to confirm our model and understand the effect of variables. First we analyze regression direct effect of all exogenous to endogenous variables. From the table below (Table 4) we can see result of direct effect which tells us the effect among variables when changing independent variable by one unit.

	Age	Education	Gender	EC Security	EC Trust	EC Legislation
E-banking	.052	.025	.069	-.061	.024	-.044
Airline	.019	.016	-.003	-.177	-.095	.112
Etail	-.040	.075	.083	-.068	.032	-.008

Table 4 Direct Impacts of Factors – AMOS Output

To understand if the relations among variables are significant we include the regression weighs from the AMOS (Table 5). Likewise, since dependent variables are of categorical type, we have performed Probit regression to understand the relation. The results of the Probit analyses can be found in Appendix B.

			Estimate	S.E.	C.R.	P
Etail	<---	ECLegislation	-0.008	0.075	-0.112	0.911
Etail	<---	ECTrust	0.032	0.082	0.385	0.7
Etail	<---	ECSecurity	-0.068	0.05	-1.377	0.168
Airline	**<---**	**ECLegislation**	**0.112**	**0.07**	**1.601**	**0.109**
Airline	<---	ECTrust	-0.095	0.077	-1.234	0.217
Airline	**<---**	**ECSecurity**	**-0.177**	**0.046**	**-3.856**	*******
Ebanking	<---	ECLegislation	-0.044	0.055	-0.801	0.423
Ebanking	**<---**	**ECSecurity**	**-0.061**	**0.036**	**-1.679**	**0.093**
Ebanking	<---	ECTrust	0.024	0.06	0.404	0.686
Etail	<---	Gender	0.083	0.087	0.954	0.34
Airline	<---	Gender	-0.003	0.081	-0.039	0.969
Ebanking	<---	Gender	0.069	0.063	1.095	0.273
Ebanking	<---	Education	0.025	0.051	0.488	0.626
Airline	<---	Age	0.019	0.096	0.198	0.843
Etail	<---	Age	-0.04	0.103	-0.389	0.697
Ebanking	<---	Age	0.052	0.075	0.694	0.488
Airline	<---	Education	0.016	0.065	0.247	0.805
Etail	<---	Education	0.075	0.07	1.075	0.282

Table 5 Regression Weights – Amos Output

Due to non-significance of some variables we can argue that results in some parts are not strong enough to prove the relation. However, the significance has to be fur-

ther analyzed to understand what it means for the purpose of the research. Johson (1999) argues that P value of significance can be lowered by increasing the sample size and therefore should be treated carefully in the research. Nevertheless, for the purpose of this research we have accepted effect of variables regardless of the significance on basis that (a) sample technique (b) sample size and (c) no prior research in area of e-commerce in Kosovo.

Compare to regression, correlation analysis do not require any priori model. Therefore, we have decided to report correlation among variables to understand the relations or associations. Correlations are reported as a number between -1 to 1. If negative correlation is reported then, one variable increases the other decreases or in case of positive correlation if one variable increase the increase as well. To accept correlation between variables the results should be significant and close to -1 or 1. For the purpose of this research we have generated correlation matrix (Table 6) from SPSS.

	Ebanking	Etail	Airline	Education	Gender	Age	EC Security	EC Trust	EC Legislation
Ebanking	1	-	-	-	-	-	-	-	-
Etail	.227[*]	1	-	-	-	-	-	-	-
Airline	.271[**]	.248[**]	1	-	-	-	-	-	-
Education	.082	.088	.033	1	-	-	-	-	-
Gender	.086	.065	-.005	-.040	1	-	-	-	-
Age	.031	-.043	.003	.136	.073	1	-	-	-
ECSecurity	-.194[*]	-.143	**-.304[**]**	-.086	.050	.151	1	-	-
ECTrust	-.101	-.019	-.062	**-.299[**]**	-.121	-.096	.192[*]	1	-
ECLegislation	-.155	-.063	-.048	**-.306[**]**	-.019	.094	**.405[**]**	**.784[**]**	1

Table 6 Correlation Matrix

According to correlation matrix (Table 6) we can see significant correlations between some variables. As mentioned above the focus is on significant correlations that close to 1 or -1. Taking this into consideration we can see a strong correlation between, ECTrust and ECLegislation r = .784, ECSecurity and ECLegislation r = .405, negative correlation Education and ECLegislation r = -306, Education and ECTrust r = -.299, Airline and ECSecurity r = .304. We also have a weak negative correlation between Ebanking and ECSecurity r = -.194 which according to the literature should

have higher correlation. However, this case has to be further discussed in the next chapter.

While there are other correlations in the matrix, we have focused in those that are relevant to the research and are closer or higher than r = +/-.300. Nevertheless, when considering correlation results we have to take into consideration, causation as well. Two variables having strong correlation with each other, does not necessarily mean that there a logical relationship or influence. Wright (1991) argues that before considering correlation analysis, researcher must have a prior knowledge about the relation of certain variables. Gardner (1999) defines causation as a variation of one independent variable to dependent variables. Therefore, for the purpose of this research the variation of independent to dependent variables are considered from regression analysis.

6 Discussion

The goal of this study is to analyze the e-commerce consumer perception and adaptation challenges in Kosovo.
Study is performed by use of Structured Equation Modeling in order to (a) test the proposed model and constructs, (b) to explore new factors and (c) to understand the influence of factors to e-commerce industry type.

The quantitative results have indicated that derived factors from literature review in case of Kosovo are moderately validated. The validated model is consisted of three factors while the proposed model had six. Change in model came after the execution of explanatory and confirmatory factor analysis. The generated model is validated and the extracted factors Legislation, Security and Trust that have the highest importance for the researcher are retained. The final generated model reported a chi-square (X2) of 48, Degree of Freedom (df) =15 and Probability level of .000. While this model has a significance level, according to Byrne (2010) this information proves that the data moderately fit to the model. However, she argues that one of the symptoms in such situation has to do with sample size. Therefore, we have considered additional indices (CFI=.801, GFI=.920 and RMSEA=.134.) to validate the results.

The final validated results are suggesting different ideas about research than originally thought. Investigating the results, it is reported from regression analysis that ECSecurity and ECLegislation are the only independent variables that have significant influence on dependent variables Airline and EBanking (Table 5). The effect of these variables is important finding and is confirmation with previous literature review from Blythe (2008) and Antoniou and Batten (2011) that security and legislation are the foundation of successful e-commerce environment. The effect of above variables leads us to another finding that Airline and Ebanking are the most used e-commerce industries in Kosovo. On the other hand, non-significant effect of independent variables on Etail is questionable. One assumption in this case is the fact of existence of only one etailer in Kosovo.

In addition to regression, we included correlation in research to identify other aspect of association among variables. Correlations confirm a strong negative associa-

tion between EC Security and Airlines. This is an indication that security negatively influences purchase of plane tickets online. While this may seem unusual and not according the existing literature review, one finding is that people in Kosovo are interested in price and therefore the security of e-commerce is not perceived too important. The next finding is the significant negative correlation between the EC-Security and Ebanking that raises some questions. However, more security implies more complexity for the usage of ebanking. Therefore, the design of e-commerce has to take into consideration security aspects as well. Yee (2004) argues that while security may affects usability of an application this situation could be overcome if both are aligned during the design phase. Similarly, ECSecurity has a low non-significant correlation with E-tail. As mentioned above, currently there is only one etailer operating in Kosovo and it is difficult to correctly estimate the impact of security in usage of etail. A study performed by Delottiee (2012) reports a slow development of etailer business in Europe which in a way it affects Kosovo as well.

Further strong and significant correlation is reported between ECLegislation and ECTrust indicating that trust of consumers to e-commerce can be positively influenced with implementation of adequate legislation framework. This confirms the suggestion from Sedith (2002) that a country that plans to develop e-commerce market should have appropriate legal framework in order to protect both providers and consumers.

The variable ECTrust is not significantly correlated to Etail, Airline or Ebanking. The measurement of ECTrust during exploratory factor analysis was based on items testing if the consumers have higher trust on local or international e-commerce provider. From results we see a non-significant effect if the provider is local or international. Therefore, a finding in this case is that consumers are not concerned if e-commerce providers are local or international. This can be considered as positive finding because according to an article published by Kohat.net, consumers in Kosovo have prejudices about the quality of local products. However, this is not true for e-commerce, which in practice creates an opportunity for future e-commerce development in Kosovo.

The factors from a priori model, ECDesign, ECSupply Chain and Providers Awareness, are dropped due to low loadings. This is an indication that consumer perception to e-commerce in Kosovo is not influenced by above dropped factors. However, this has to be further analyzed.

7   Conclusions

This paper attemptes to clarify the situation of e-commerce in Kosovo by analyzing the perception of consumers and identifying adoption challenges. The research was conducted by reviewing literature, use of quantitative methods and exploration of factors that were statistically validated. The important factors that were explored during the research are E-commerce Security, E-Commerce Legislation and E-Commerce Trust. These factors are then used to measure the consumer perception and

adoption of the e-commerce in area of ebanking, etailer and purchase of plane e-tickets.

The findings of the research conclude that ebanking is still the most used e-commerce service in Kosovo followed by purchase of airline tickets and etailers. One conclusion of low usage of etailers is low number of actual etailers in Kosovo i.e. only one. Further, it can be concluded that security has an influence in purchase of airline tickets, but not necessarily in usage of ebanking or etail. Legislation plays a vital part of developing and regulating e-commerce market in Kosovo and the government should put their focus in this area. However, Legislation does not influence the purchase of airline tickets, or the usage of ebanking and etail. Same applies with trust, but in this case it has been concluded that consumers in Kosovo equally trust local and international e-commerce providers. In addition, since consumers are not influenced by security, legislation or trust, then we can conclude that consumers in Kosovo are price driven. Moreover, since e-commerce in Kosovo is in early stage of development and with only ebanking offered by banks, airplane tickets offered by international providers and one local etailer, we can conclude that market of e-commerce is still immature. In order to increase the maturity of the market, public and private institutions must work closely to develop a sustainable e-commerce environment.

7.1   Limitations and future direction for research

The paper has critically analyzed the e-commerce in Kosovo, however limitations in some areas are identified as following:

Sampling Technique – The employment of convenient sampling method has created a sample that is not fully representative compare to the rest of population. The sample is mainly consisted of responders from bank with (a) higher income than average (b) higher education, (c) majority with Internet connection and (d) hold bankcard.

Sample size – the sample size for the research was N=123. According, to the study from Costello and Oseborn (2010) the rule of thumb for exploratory facto analysis is the ratio 10x1 i.e. ten responses per variable. For this research we have reached the ratio of 4.7x1 i.e. 123/26. Limitation of sample caused some regression and correlation to be non-significant.

Previous research – No previous research in area of e-commerce was performed in Kosovo. Therefore, the literature review is relaying on studies from EU countries that far more advanced in technology and infrastructure.

In general, to have a better generalization, future researches should take into consideration above limitations. In addition, future researches should analyze in more details removed factors (EC-Design, Providers Awareness and EC-Supply Chain). Besides these factors, future research should critically analyze the pricing aspect and its influence on e-commerce consumers. Overall as mentioned in conclusion section, since the e-commerce market in Kosovo is still immature the research should be repeated after two years with above suggestions in mind.

References

Al-Ma'aitah, M. and Shatat, A. (2011) 'Empirical Study in the Security of Electronic Payment Systems', *International Journal of Computer Science Issues,* 8(4), pp. 393-401.

Al-Slamy, N. M. A. (2008) 'E-Commerce security', *International Journal of Computer Science and Network Security,* 8(5), pp. 340-344.

Antoniou, G. and Batten, L. (2011) 'E-commerce: protecting purchaser privacy to enforce trust', *Electron Commer Res,* 11, pp. 421-456.

Barber, W. and Badre, A. (1998) *Culturability: The Merging of Culture and Usability,* Available at: *http://research.microsoft.com/en-us/um/people/marycz/hfweb98/barber/* (Accessed: 15th September 2013).

Bargh, A.J. and Katelyn Y. A. McKenna, KY.A. (2004) ' The Internet and Social Life', *Annual Review of Psychology,* 55, pp. 573-590.

Barsauskas, P., Sarapovas, T. and Cvilikas, A., (2007) 'The evaluation of e-commerce impact on business efficiency', *Baltic Journal of Management,* 3(1), pp. 71-91.

Belkhamza, Z. (2009) 'The Effect of Perceived Risk on the Intention to Use E-commerce: The Case of Algeria', *Journal of Internet Banking and Commerce,* 14(1), pp. 2-10.

Being (2014) *Heywood cases (negative error variances or squared multiple correlation that is greater than 1),* Available at: *http://zencaroline.blogspot.com/2007/05/heywood-cases-negative-error-variances.html* (Accessed: 26th December 2013).

Bentler, P.M. (1990), "Comparative Fit Indexes in Structural Models," *Psychological Bulletin*, 107 (2), 238-46.

Bergman,J.(2010) *Kosovo's Growing Pains - TIME,* Available at:*http://content.time.com/time/world/article/0,8599,1964625,00.html* (Accessed: 23rd December 2013).

Blythe, S. (2008) 'Croatian computer laws: promotion of growth in E-commerce via greater cyber-security', *European Journal Law of Economy,* 26, pp. 75-103.

Bin, Q., Chen, S.-J. and Sun, S.Q. (2003) 'Cultural differences in e-commerce: a comparison between the U.S. and China', *Journal of Global Information Management,* 11, pp. 48-55.

Brown, T. (2006) *Confirmatory Factor Analysis for Applied Research (Methodology in the Social Sciences)*, New York, NY: The Guilford Press.

Carraher, S. M., and Buckley, M. R. (1991) 'The effect of retention rule on the number of components retained: The case of the Pay Satisfaction Questionnaire',*Proceedings of the Southern Management Association.*
Chau, P.Y.K., Cole, M., Massey, A.P., Montoya-Weiss, M. and Keefe, R.M., (2002) 'Cultural differences in the online behavior of consumers', *Communications of the ACM,* 45, pp. 138-143.

Cortina J. (1993) 'What is coefficient alpha: an examination of theory and applications, *Journal of applied psychology,* 78, pp. 98-104.

Costello, B. A. and Osborne, J. W. (2005) 'Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis', *Practical Assessment Research & Evaluation*, 10(7), pp. 2-9.

Daft R. and Lengel R. (1986) 'Organizational Informational Requriments, media Richness and Structural Design', *Management Science,* 32(5), pp. 554-571.

Dean, D. (2012) *The Internet Economy in the G-20*, Boston: The Boston Consulting Group.

Deloittee (2012) *European eCommerce Assessment 2012 Benchmarking the top 200 in online retail*, The Netherlands: Deloittee.

De Mooij, M. (2004) *Consumer Behavior and Culture: Consequences for Global Marketing and Advertising*, Thousand Oaks, CA: Sage.
Last name, Initials (Year) *Book title,* Publisher, Town (10 pt)

Earley, C.P. and Erez, M. (1993) *Culture, Self-Identity, and Work*, USA: Oxford University Press.

Efendioglu, A.M. and Yip, V.F. (2004) 'Chinese culture and e-commerce: an exploratory study', *Interacting with Computers,* 16, pp. 45-62.

Farrell, H. (2003) 'Constructing the International Foundations of E-Commerce: The EU-U.S. Safe Harbor Arrangement', *International Organization,* 57(2), pp. 277-306.

Fazliu A. (2012) *Internet Penetration and usage in Kosovo*, Prishtina: The Kosovo Association of Information and Communication Technology.

Ferber, R. (1977) 'Research by Convenience', *Journal of Consumer Research,* 4(1), pp. 57-58.

Gardner, R., G. (1999) 'Correlation, Causation, Motivation, and Second Language Acquisition', *Department of Psychology University of Western Ontario,* 41(1), pp. 11-23

Gefen and Straub (2000) 'The Relative Importance of Perceived Ease-of-Use in IS Adoption: A Study of e-Commerce Adoption', *Journal of AIS,* 1, pp. 1-30.

Hansen J. V. (2001) 'Internet Commerce Security: Issues and Models for Control Checking', *The Journal of the Operational Research Society,* 52(10), pp. 1159-1164.

Henderson, J., Dooley, F. and Akridge, J. (2004) 'Internet and E-Commerce Adoption by Agricultural Input Firms', Review of Agricultural Economics, 26(4), pp. 505-520.
Hitt, L. and Brynjolfsson, E. (1996) 'Paradox Lost? Firm-Level Evidence on the Returns to Information Systems Spending', *Management Science,* 42(4), pp. 541-558.

Hsieh, M. (2012) *Ecommerce: The Supply Chain Perspective,* Available at: *http://pando.com/2012/07/02/ecommerce-the-supply-chain-perspective/* (Accessed: 15th December 2013).

Kalakota, R. and Whinston A.B. (1997) *Electronic Commerce: A Manager's Guide*, Reading: Addison-Wesley.

Kendall, J., Tung, LL, Chua, KH. NG. CH. and Tan, SM (2001) 'Electronic commerce adoption by SMEs in Singapore', *In Proceedings of the 34th Annual Hawaii International Conference on System Sciences,* 34, pp. 10.

Kolenikov, S. and Bollen, K.A.(2007) 'Testing Negative Error Variances: Is a Haywood Case a Symptom of Misspecification?' *Unpublished paper available at: http://web.missouri.edu/~kolenikovs/papers/heywood-4.pdf*

Kosovo Bankers Association (2012) *The Kosovo Banker*, Prishtina: Kosovo Bankers Association.

Koufaris, M. (2002) 'Applying the technology acceptance model and flow theory to online consumer behavior', *Information Systems Research,* 13(2), pp. 205-223.

Kshetri, N.B. (2001) 'Determinants of the locus of global e-commerce. *Electronic Markets*', 11, pp. 250-257.

Lee, H.L. and Billington, C. (1993) 'Material Management in Decentralized Supply Chains ', *Operations Research,* 41(5), pp. 835 - 847.

Lee, Ch., Sh. (2001) 'An analytical framework for evaluating e-commerce business models and strategies', *Internet Research,* 11(4), pp. 349-359.

Lewis, Ira. (2001) 'Logistics and Electronic Commerce: An Interorganizational Systems Perspective', *Transportation Journal,* 40(4), pp. 5-13.

Li, M. and Ye, M. (1999) 'Information technology and firm performance: Linking with environmental, strategic and managerial contexts', *Information & Management,* 35, pp. 43-51.

Liao, H.,Proctor, R.W. and Salvendy, G. (2008) 'Content preparation for cross-cultural e-commerce: a review and a model', *Behaviour & Information Technology,* 27(1), pp. 43-61.

Lightner N. J., Yenisey M. M., Ozok A. A., Salvendy G., (2002) 'Shopping behavior and preferences in e-commerce of Turkish and American university students: implications from cross-cultural design', *Behavior and Information Technology,* 6(21), pp. 373-385.

Lo, B. and Gong, P. (2005) 'Cultural Impact on the Design of E-Commerce Websites: Part I Site Format and Layout', *Issues in Information Systems,* 6(2), pp. 182-189.

McCort, D.J. and Malhotra, N.K., (1993) 'Culture and consumer behavior: toward an understanding of cross-cultural consumer behavior in international marketing.' *Journal of International Consumer Marketing,* 6, pp. 91-127.

Mentzer, J.T., DeWitt, W., Keebler, J.S., Min, S., Nix, N.W., Smith, C.S. and Zacharia, Z.G (2001) 'Defining Supply Chain Management', *Journal of Business Logistics,* 22(2), pp. 1-25.

Molla, A., Heeks, R. and Balcells, I. (2009) 'Adding clicks to bricks: a case study of e-commerce adoption by a Catalan small retailer', *European Journal of Information Systems ,* 16, pp. 424-438.

Ngai, E.W.T. and Wat, F.K.T. (2004) 'A literature review and classification of electronic commerce research', *Information and Management,* 2002(39), pp. 415-429.

Nielsen, J. (2005) *Foreword to book. In Usability and Internationalization of Information Technology*, Mahwah, NJ: Lawrence Erlbaum Associates.
OECD Economic Outlook (2000) *E-commerce: impacts and policy challenges*, Paris: OECD

Oxley, J. E., and Yeung. B. (2001) 'E-commerce readiness: institutional environment and international competitiveness', *Journal of International Business Studies,* 32(4), pp. 705-723.

Pavlou, P. A., and Fygenson, M. (2006) 'Understanding and predicting electronic commerce adoption: An extension of the theory of planned behavior', *MIS Quarterly,* 30, pp. 115-143.

Pearson, H. (2000) 'E-Commerce Legislation: Recent European Community Developments', *Journal of Internet Law,* 4(2), pp. 1-8.

Petrovic, D. and Kovacevic, I. (2012) 'Distrust as Obstacle to e-Commerce Development in Serbia', *Journal of Theory and Practice Management,* 65, pp. 71-77.

Persefoni, P., Stavros, V., Giannoula, F. and Anastasios, K. (2001) 'Comparison of e-commerce Behaviour among women consumers of Balkan and Baltic regions', *Economic Sciences,* 10(2), pp. 1-14.

Rahman, S.M. and Bignall, R.J. (2001) *Internet Commerce and Software Agents: Cases, Technologies, and Opportunities*, Hershey: Idea Group Publishing.

Reise, S. P., Widaman, K. F., and Pugh, R. H. (1993) 'Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance', *Psychological Bulletin,* 40(1), pp. 115–148.

Russell, J.L. and Russell, L. (2006) *Change Basics*, Baltimore: ASTD Press.

Steenkamp, Jan-Benedict E.M., Batra, R. and Dana L. (2003) 'How Perceived Brand Globalness Creates Brand Value', *Journal of International Business Studies,* 1(34), pp. 53-65

Trivun, V. and Siljazdic, V. (2007) 'Harmonization of B&H Legislation with EU Law in domain of E-Legislation', *International Conferences on Economics and Management of Networks - EMNet,* , pp. 1-15.

United Nations Commission on International Trade Law (2013) *UNCITRAL Model Law on Electronic Signatures,* Available at: *http://www.uncitral.org/uncitral/en/uncitral_texts/electronic_commerce/2001Model_signatures.html* (Accessed: 18th August 2013).

Wright, S. (1991) 'Correlation and Causation', *Journal of Agriculture Research,* 20(7), pp. 557-585.

Yee, K (2004) 'Aligning security and usability', *Security & Privacy, IEEE ,* 2(5), pp. 48-55.

Appendix A: Survey Questioner

Q1	Companies do not invest in e-commerce because they lack technical know how
Q2	Companies that invest in IT are more profitable than those who don't
Q3	Companies do not invest in e-commerce because cost of implementation is very high
Q4	Lack of technical know-how in companies is due to low investment in technology trainings/education
Q5	The e-commerce website must have the necessary security measures to prevent hackers from accessing my information
Q6	Money transaction in the Internet are secure (Polychronidou et. al., 2010)
Q7	Personal information stored in e-commerce are not secure
Q8	E-commerce must provide adequate information about the liabilities in case of incident (money stolen, product not delivered)
Q9	Government is not doing enough to promote e-commerce legislation
Q10	E-Commerce providers must do more promote legislations in Kosovo
Q11	Digital signatures are equivalent to l hand written signatures
Q12	Adapting e-commerce legislation that protects consumer right, will increase consumer confidence
Q13	E-commerce in Kosovo is challenged with delivery of products purchased online because of missing address system
Q14	Consumers in Kosovo don't purchase online because they are afraid that product will not be delivered as expected
Q15	I'd use e-commerce only if the shop is far away
Q16	E-commerce would not let me feel the product and consult shop assistant for advice
Q17	Kosovo consumers trust cash payments more than electronic payments (card, e-banking etc)
Q18	Design of e-commerce is important to increase consumers trust
Q19	Local e-commerce providers are more trustworthy than international
Q20	International e-commerce providers are more trustworthy than local
Q21	The e-commerce website in Kosovo are not well designed and not user friendly
Q22	Banks in Kosovo do not process online transaction (over the Internet) for e-commerce businesses
Q23	Using e-commerce would not let me use the opportunity to negotiate prices in store

Q24	I would use e-commerce than go to shop if the price of the product is lower than in the store
Q25	I would use e-commerce than go to shop if different price options are avalialbe (bulk purchase etc.)
Q26	Colors increase credibility of e-commerce providers (website)
Q27	Which of these colors would you prefer in a e-commerce website
Q28	Do you have Internet connection at home
Q29	Do you use e-banking
Q30	Do buy/purchase products (books, music, clothes etc.)
Q31	Do you buy plane tickets online
Q32	Do you trade stocks online
Q33	Do you participate in online auctions (example ebay)

	Demographics
Q34	What is your education level
Q35	Gender
Q36	Do you hold a credit or debit card?
Q37	Age

Appendix B: SPSS Result – Probit Analysis

Dependent Variable: **Ebanking**

	Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
PROBIT[a]	Trust	.000	.068	.004	.997	-.132	.133
	ECSS	-.024	.041	-.585	.559	-.104	.056
	Legislation	-.014	.063	-.228	.820	-.137	.108
	Intercept	-2.373	.123	-19.327	.000	-2.496	-2.251

a. PROBIT model: PROBIT(p) = Intercept + BX

Dependent Variable: **Etail**

	Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
PROBIT[a]	Trust	.010	.075	.127	.899	-.138	.157
	ECSS	-.035	.046	-.755	.450	-.125	.056
	Legislation	-.008	.069	-.119	.905	-.144	.127
	Intercept	-2.476	.137	-18.134	.000	-2.613	-2.339

a. PROBIT model: PROBIT(p) = Intercept + BX

Dependent Variable: **Airline**

	Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
PROBIT[a]	Trust	-.051	.074	-.687	.492	-.195	.094
	ECSS	-.092	.046	-1.996	.046	-.183	-.002
	Legislation	.059	.068	.866	.387	-.074	.191
	Intercept	-2.332	.131	-17.804	.000	-2.463	-2.201

a. PROBIT model: PROBIT(p) = Intercept + BX

Motives for Cultural Consumption:
Developing a new scale

Maria Manolika[1] and Alexandros Baltzis[1]

[1] Department of Journalism & Mass Communications, Aristotle University of Thessaloniki
Email: mmanolik@jour.auth.gr; maria.manolika@yahoo.gr

Abstract. Aim of this paper is to present the background and methodology of a research designed to investigate consumer motivation for cultural goods (e.g. recorded music, books, museums, etc.). In the first part, the development of cultural consumption is assessed and the significance of understanding and evaluating the cultural consumer motivation is outlined. The second part of the paper includes a systematic presentation of past research and theoretical approaches, concluding that there is a lack of a multi-dimensional consumer motivation scale, specifically designed for the field of culture. The third part outlines the proposed methodology and describes the dimensions of a new multi-dimensional scale to trace and evaluate the cultural consumer motivation, measuring three main factors: emotional, cognitive, and social. Finally, the paper presents the main features of the planned survey and analyses the significance of this research for understanding how social and demographic factors affect the motivation of the cultural consumer.

Keywords: *consumer motivation, cultural industries, cultural consumption*

1 Introduction: Rapid growth and chances in public demand

During the last decades, cultural organizations and industries (e.g. museums, theaters and art galleries) have increased in number and size worldwide, and have become one of the fastest growing sectors of the leisure industry. The enhancement of the domestic culture and local history, the provision of various recreational opportunities, and the contribution to the domestic economy by stimulating tourism businesses, include the main advantages deriving from this rapid growth (Getz, 2008; Long and Perdue, 1990). Consequently, cultural industries and organizations have significant economic, socio-cultural and political impacts on the regions they operate – a fact recognized by local governments, as well as by supranational and transnational organizations and governing bodies, like the OECD, the UN, and the EU.

Since competition among cultural organizations and tourist destinations is increasing in an increasingly privatized and growing creative economy, the need for further information, specifically the analysis of motivations for attending cultural organizations and consuming cultural goods, has become crucial for policy makers

and managers (Getz, 2008; UNESCO/UNDP, 2013). According to Crompton and McKay (1997), there are three major reasons for identifying and understanding the cultural consumer motivation.

- Motives play a major role in designing and offering suitable products for consumers, who seek to satisfy a variety of divergent needs. As a result, identification of these needs is a prerequisite for effective development of suitable products.
- Secondly, motivation is an antecedent of satisfaction and satisfaction is a precursor to repeat visits. Most cultural organizations are highly dependent on repeat visitors. If needs are fulfilled, satisfaction and return visits will come as a result. Therefore, organizers and managers will be informed about the needs which visitors seek to satisfy.
- Finally, identifying and prioritizing motives is a key element in understanding a consumer's decision process. Accordingly, marketing endeavours and activities will become more effective and efficient.

Therefore, organizers and managers are likely to assert that their fundamental goal is to provide high quality products and satisfying experiences that people perceive to be a good value in order to increase the probability visitors to return in the future or recommend the organization to others (Lee et al, 2007). Albeit researchers have examined commonly targeted topics (e.g. socio-economic impact, marketing strategies, and management), little research has been conducted on the consumers motivation. A shift of focus from the organizers and cultural goods to the public and its special needs is required. In order to cover these deficiencies, the present study focuses on the motives of cultural consumption. The main objective of this paper is to critically approach the most relevant theoretical frameworks used to understand individuals' motives of cultural consumption and outline the steps for developing a multi-dimensional motivation scale to better account for the consumer–product relationship.

2 Literature Review: Motives for cultural consumption and theoretical approaches

A motive can be described as an internal factor that arouses, directs and integrates a person's behaviour and activity. As a result, motives are the driving force behind all human behaviours. In accordance with basic motivation theory, when a need, a want or a goal occurs, a dynamic process of internal psychological factors appears, that generate an uncomfortable level of tension within individuals' minds and bodies (Alderfer, 1989). These inner needs and the resulting tension lead to actions designed to release tension, by satisfying these needs (Maslow, 1954; Murray, 1938). From a marketing perspective, products can be designed and marketed as solutions to consumer's needs. Thus, the consumer will be motivated to buy, only if he or she perceives the purchase of a product or a service as having a positive effect on satisfying a conscious or unconscious need (Mill and Morrison, 1985).

Although motives are the starting points that trigger decision processes and have an impact on the selective direction of behaviour (the choice of one action over another), the energisation (the amount of the required energy or effort), and the persistence of the acts (their persistence in time), most research concerning cultural consumption neglects to study in detail the crucial role of motives in people's behaviour. For example, most studies have focused mainly on demographic or social factors (Hooper-Greenhill, 1994), on specific and limited range of cultural goods, such as museums or art galleries (Prentice et al, 1997; Slater, 2007) and on a restricted range of motives and theoretical approaches (Reynolds, 2007).

As there has not been an extensive research on the motivations of cultural consumption and the theoretical background is also limited, it is pertinent to draw attention on leisure and event motivation theory. In this field, three theoretical frameworks have been developed: Maslow's (1945) human need hierarchy, Iso-Ahola's (1980; 1982) escape-seeking dichotomy and Dann's (1981) push and pull factors. Crompton (1979), based on push-pull factors (Dann, 1981), investigated the motives for pleasure-oriented tourism. He identified seven push and two pull factors, including: escape, exploration of self, relaxation, prestige, regression, kinship enhancement, social interaction, novelty and education. Also, Uysal et al (1993) using the seeking-escaping theory (Mannell and Iso-Ahola, 1987), identified five motivational factors of audiences at a community festival in South Carolina.

However, it can be concluded that until nowadays, most festival motivation studies are based on the above mentioned theoretical approaches, utilizing a pool of items that ranges from 19 to 34 and generate about four to six dimensions after various factor analyses (Uysal et al, 1993). Most studies (Lee et al, 2004) used factor analysis to reduce the size of items. The extracted factors served as a foundation for subsequent analyses – such as segmentation analysis – to identify the various groups of participants with homogenous preferences (see: **Table 1**).

Table 1. Review of selected studies on festival and event motivation

Researchers	Dimensions of motives	Event studied	Scale used
Uysal et al (1993)	Escape; excitement/thrills event novelty; socialization; relaxation	Corn Festival, South Carolina, USA	24 statements 5-point Likert Scale
Mohr et al (1993)	Socialization; escape; family togetherness; excitement/uniqueness; event novelty	Freedom Weekend Aloft, South Carolina, USA	23 motive items 5-point Likert Scale
Crompton and Mckay (1997)	Cultural exploration; novelty/regression; gregariousness; recover equilibrium; known-group socialization; external interaction/socialization	Fiesta in San Antonio, Texas, USA	31 motive items 5-point Likert Scale
Lee et al (2004)	Cultural exploration; family togetherness; novelty; escape event attractions; socialization	2000 Kyongju World Cultural Expo., South Korea	34 motive items 5-point Likert Scale

Researchers	Dimensions of motives	Event studied	Scale used
Yuan et al (2005)	Festival and escape; socialization; wine; family togetherness	Vintage Indiana Wine and Food Festival, Indiana USA	27 motive items 7-point Likert Scale
Dood et al (2006)	Socialization; family togetherness; excitement/uniqueness; escape; event novelty	GrapeFest Festival, Texas Vintage Indiana Wine and Food Festival, Indiana USA	14 motive items 7-point Likert Scale

Note: Partly adapted from Lee et al (2004)

According to the literature review (Lee et al, 2004), we can see that there are five main motivations that appear in most studies:

- Socialization (a need of interacting with other people),
- Family togetherness (a desire to tie family bonds),
- Escape (a change from routine life and recovery),
- Novelty (a need for variety of new things and a desire of obtaining an unusual experience, not available in the routine life),
- Cultural exploration (a need for learning new information and exploring new and different cultures).

The above mentioned motives explain to a limited extent the reasons that people visit festivals and events and until now they have not been examined in the field of cultural consumption. In addition, research on the motivation including a wider range of sectors of cultural consumption still remains limited.

3   Proposed Methodology

Proposed hierarchical structure of cultural motives

Despite the fact that five main motivations appear in the festival and leisure literature, it seems that they cannot fully capture cultural consumers' motivation. In addition, they are not distinctive enough and they overlap each other. For example, cultural exploration combines both a need for learning and exploring new cultures in the same facet (Lee et al, 2004), when in fact they should be distinct. As a result, the present study proposes a hierarchical structure of consumers' motivation.

In particular, there are nine first order motivational factors that explain cultural consumption behavior. In turn, these factors organized into three higher order motivational factors, namely Emotional (enjoyment/entertainment, emotional arousal, escape/recovery and novelty/curiosity), Cognitive (cultural exploration, learning and self-actualization) and Social (external-internal socialization and identification-discrimination). This conceptual framework was drawn on the leisure and tourism literature (Uysal et al, 1993) and on motivation theories (Deci and Ryan, 1985; Maslow, 1954), as motivations of cultural consumption have not previously been measured by multiple items. The emotional factor refers to the human need for

experiencing feelings of joy and satisfaction, rest and relaxation and seeking new and novel experiences. The cognitive factor refers to the peoples' need to have an opportunity for learning, expanding their knowledge, or an urge to do something worthwhile in general. Finally, the social factor concerns the need of people to be with others, to meet new friends and gain a sense of belonging.

Steps for development of the research instrument

Initially, a pool of candidate items will be developed based both on theoretical background and previous instruments (sources: see **Table 1**). Prior to the formal survey, a pilot study will be conducted to assess the clarity of the items and estimate their internal consistency. Students from the Aristotle University of Thessaloniki will participate in the pilot study. Exploratory and confirmatory factor analysis will be conducted to understand the underlying structure of the new scale. Finally, test-retest procedures will be used to examine scale's temporal stability.

Final questionnaire

A self-administered questionnaire will be developed as the instrument for collecting data. It will include two sections. The first section will be designed to assess the visitors' motivation on a 5-point Likert scale, ranging from "strongly agree" (5) to "strongly disagree" (1). The second section will include socio-demographic data about the respondents, including age, gender, and education, region of residence, occupation, and income level.

Site and population of the study

The final questionnaire will be distributed to different cultural organizations in Thessaloniki – a major Greek urban centre – to explore the consumers' motivation both for tangible and intangible cultural goods. Museums, theatres, musical performances (intangible cultural goods) and music stores, book stores and art galleries (tangible goods) will be the targeted organizations. The population for the study will be the attendees for each cultural organization.

4   Conclusion

Up to now, in the field of cultural consumption, there has not been developed a motivational scale that includes a wide array of potential consumer motives. Therefore, the main objective of the current research is to fill this gap, by developing and validating a multi-dimensional scale that takes into account the different motivations of the cultural consumers. Indeed, a multi-dimensional consumer motivation scale is useful, as it may have more predictive power for cultural buying behavior because it takes into account a wider range of consumer motivations. In a

theoretical level, this research will increase the present theoretical background and will shed some light to aspects that have not been explored yet.

Moreover, this research has also practical relevance for marketing managers who are confronted with the difficulty of creating and launching successful cultural goods. Firstly, the different motives may provide an interesting line of reasoning for organizers and be a useful development tool for policy makers too. The public of culture that is differently motivated should be targeted with effective marketing communications emphasizing and triggering specific motivations. For example, communications to cognitively motivated consumers should focus on the pleasure of acquiring new knowledge. Segmenting the public and understanding its features based on motivations will be important for successful marketing strategies in the future.

However, there are also some limitations. The present thesis explores for the first time the motives for cultural consumption and tests a scale for measuring and evaluating consumer motivation. Future research should expand in exploring the motives for cultural consumption in other locations, as well as in different countries and cultures. Only then, results can be generalized and the theory of cultural consumption can be further elaborated. Additionally, motives are important prerequisite for attitudinal and behavioral change (Moorman and Matulich, 1993), and produce more stable and enduring attitudes (Batra and Michael, 1986; Celsi and Olson, 1988; Hoch and Deighton, 1989). Therefore, the relationship between motivation-attitude and behavioral intentions should be in the focus of future research.

References

Alderfer, C. P. (1972) *Existence, relatedness, and growth: Human needs in organizational settings*. Free Press, New York.

Batra, R. and Michael, R. (1986) Situational effects of advertising repetition: the moderating influences of motivation, and opportunity to respond. *Journal of Consumer Research*, 12(4), p. 432-45.

Celsi, R. L. and Olson, J. C. (1988) The role of involvement in attention and comprehension processes. *Journal of Consumer Research*, 15(2), p. 210-24.

Crompton, J. L. and McKay, S. L. (1997) Motives of visitors attending festival events. *Annals of Tourism Research*, 24(2), p. 425-39.

Dann, G. (1981) Tourism motivations: Appraisals. *Annals of Tourism Research*, 8(2), p. 189-19.

Deci, E. L. and Ryan, R. M. (1985) *Intrinsic motivation and self-determination in human behavior*. New York, Plenum Press.

Dood, T., Yuan, J., Adams, C. and Kolyesnikova, N. (2006) Motivations of young people for visiting wine festivals. *Event Management*, 10(1), p. 23-33.

Getz, D. (2008) Event tourism: definition, evolution, and research. *Tourism Management*, 29(3), p. 403-28.

Hoch, S. J. and Deighton, J. (1989) Managing what consumers learn from experience. *Journal of Marketing*, 53(2), p. 1-20.

Hooper-Greenhill, E. (1994) *Museums and their visitors*. London/New York, Routledge.

Iso-Ahola, S. E. (1980) *The social psychology of leisure and recreation*. Dubuque IA, William C. Brown Company.

Iso-Ahola, E. (1982) Towards a social psychology theory of tourism motivation: a rejoinder. *Annals of Tourism Research*. 9(2), p. 256-62.

Lee, C. K., Lee, Y. K. and Wicks, B. E. (2004) Segmentation of festival motivation by nationality and satisfaction. *Tourism Management*, 25(1), p. 61-70.

Lee, S. Y., Petrick, J. F. and Crompton, J. (2007) The roles of quality and intermediary constructs in determining festival attendees' behavioral intention. *Journal of Travel Research*, 45(4), p. 402-12.

Long, P. T. and Purdue, R. R. (1990) The economic impact of rural festivals and special events: assessing the spatial distribution of expenditures. *Journal of Travel Research*, 28(4), p. 10-4.

Mannel, R. and Iso-Ahola, S. (1987) Psychological nature of leisure and tourism experience. *Annals of Tourism Research*, 14(3), p. 314-29.

Maslow, A. H. (1954) *Motivation and personality*. New York, Harper & Row.

Mohr, K., Backman, K. F., Gahan, L. W. and Backman, S. J. (1993) An investigation of festival motivations and event satisfaction by visitor type. *Festival Management and Event Tourism*, 1(3), p. 89-97.

Moorman, C. and Matulich, E. (1993) A model of consumer preventive health behaviors: the role of health motivation and health ability. *Journal of Consumer Research*, 20(2), p. 208-28.

Murray, H. A. (1938) *Explorations in personality: A clinical and experimental study of fifty men of college age*. New York, Oxford University Press.

Prentice, R. C., Davies, A. and Beeho, A. (1997) Seeking generic motivations for visiting and not visiting museums and like cultural attractions. *Museum Management and Curatorship*, 16(1), p. 45-70.

Reynolds, A. (2008) Social identity as a motivator in cultural festivals. *Visitor Studies*, 11(1), p. 3-15.

Slater, A. (2007) Escaping to the gallery: understanding the motivations of visitors to galleries. *International Journal of Nonprofit and Voluntary Sector Marketing*, 12(2), p. 149-62.

UNESCO/UNDP (2013) *Creative economy report 2013 special edition. Widening local development pathways*. New York, UNDP.

Uysal, M., Gahan, L. and Martin, B. (1993) An examination of event motivations: A case study. *Festival Management and Event Tourism*, 1(1), p. 5-10.

Yan, J., Cai, L., Morrison, A. and Linton, S. (2005) An analysis of wine festival attendees' motivations: a synergy of wine, travel and special events. *Journal of Vacation Marketing*, 11(1), p. 41-58.

Information and Communication Technologies

Sign Language Recognition using Kinect

Edon Mustafa[1], Konstantinos Dimopoulos[2]

[1]South-East European Research Centre, University of Sheffield, Thessaloniki, Greece
[2]CITY College- International Faculty of the University of Sheffield, Thessaloniki, Greece
edmustafa@seerc.org, k.dimopoulos@city.academic.gr

Abstract. Sign language is used for communication among hearing impaired people. Their communication is hardly understood by normal hearing people. To facilitate communication between these two groups this paper presents a system that uses Microsoft Kinect device to build a translation system that translates signs from sign language to spoken language. The developed system uses SigmaNIL framework to provide hand shape recognition which is not supported by official and other SDKs for Kinect. The resulting system is capable of translating a limited vocabulary of Kosova Sign Language. The system has been tested with native and non-native speakers of sign language and in general achieves high accuracy varying by components of sign language.

Keywords: Sign Language Recognition Kinect Hand Shape

1 Introduction

Sign language is used by the community of Hearing Impaired (HI) people as the main means of communication. Compared to Normal Hearing (NH) people that use oral communication, HI people use visual signs that involve the use of both hands, the head, facial expressions and body postures. Exchange of information between a HI and a NH person is difficult since the NH person must be able to understand the sign language of the HI person. Because of the impairments that HI people have on hearing, the responsibility of learning the sign language falls on the NH person or alternatively a third person that understands both oral and sign languages and would do the necessary translation from one language to the other. A third alternative, is to use a computer to act as a translator. Ideally the computer system should watch HI person performing sign language and then translate it to speech for NH person; to accommodate of a dialogue, the system should also listen the NH person speaking and should translate it into a form that the HI person could understand; that is either sign language or text. Such a system has been developed here, focusing on the first translation direction: from signs to speech, since it is challenging and not developed at levels of being used in real world scenarios. The developed system employs the Microsoft Kinect device to infer necessary features for sign description. The Kinect

consist of various sensors to supports voice, movement and gesture recognition [1]. The skeleton tracking feature that is made possible through its depth sensor is used in this system [2]. In order to program the Kinect a number of SDKs exist, which offer skeletal tracking but not hand shape tracking (thus not allowing for finger level precision, which can be important for this application), with one exception: SigmaNIL. SigmaNIL wraps around other SDKs and offers the hand shape recognition feature [3]. Programming with SigmaNIL is challenging since it is on beta stage. It lacks the documentation, support and some functionalities like hand shape tracking of both hands and skeleton positions of other body parts. Solutions for these functionalities have been developed and they are presented at later section. The resulting system is capable of recognizing alphabet letters, digit numbers, words and sentences from the Kosova Sign Language (KSL). The system was tested by native speakers of KSL with recognition accuracy that depends on the component being recognized and depends on whether one hand or two were involved in sign making, whether they occluded each other and whether they included movement or not.

2   Background Research

2.1   Sign Language

Hearing impairments can vary from limited hearing to complete deafness. Since the process of learning how to speak involves the use of auditory feedback, people who are born with hearing impairments also face difficulties with speech even though they may have nothing physically wrong with their vocal system. Communication between two people with hearing impairments involves the use of hands to describe the shape of something, or to describe actions [4]. For example it is common (also to people with no hearing impairments) to express the action "go from this place to that place", by pointing first to the source and "drawing" a line toward the destination. Communication is augmented with the simultaneous use of facial expressions and body postures together with hands gestures to fully express in this language any meaning [4]. There are many situations where hand gestures are used in place of oral communication. Babies may use simple gestures to express their needs before they learn to speak or adults in situations where speech is impossible or not appropriate. However, in contrast to these situations sign language is structured and has rules for composition and interpretation [4]. Sign language is taught to children with hearing impairments as a mother language [5], but it is a not universal language since as each vocally spoken language has its own sign language dialect [6]. This means for example that Greek sign language is different from English sign language. Specifically, a sign language mimics the fundamental properties that are present in the spoken language, like the grammar and vocabulary and it is used to express complex as well as abstract meanings [7]. Sign language is classified as a natural language and in many countries it is legally recognized.

Sign making involves the upper part of human's body, where parts of upper body are categorized as being manual or non-manual (automatic) features. Manual features involve hands [4] while non-manual features involves facial expressions (eye blinking, mouth shapes etc.), body postures and head movements [7]. Because manual features express most of the meaning in sign language and because of the limited time of the project the developed system addresses only manual features (those related to hands) while non-manual features although important for expressing grammatical features [8] they are not addressed in this system. Comparing sign languages with vocally expressed languages, there is a significant difference: signs (being visual in their nature) allow for simultaneous (parallel) processing of information, contrary to vocally expressed languages where only one sound can be perceived at a time, and thus are linearly processed [8]. However the amount of information conveyed in a given interval is the same with both languages [9]. Sign languages do not use articles (e.g. "*the*"), conjunctions (e.g. "*and*") and copulas (e.g. "It *is* raining") [8]. However there are also many vocally expressed languages that miss these features. Non-manual features are used for showing grammatical features like making questions, negatives [8] and showing boundaries between sentences [7]. Yes and no questions are signaled by raising the eyebrows, negatives by shaking the head and mouth shapes to signal degree (intensity) of something being communicated [7]. For example in American Sign Language to express the sentences "it is raining" and "it is *not* raining", the same gestures are used with the difference that in the second case, the head is horizontally rotated to a negation. The sign location is used to indicate tense (present, future, past). Signs performed near body refer to present, signs in front of the body to the future and signs behind the shoulders indicate something in the past [10]. Sign language is a natural language, very rich in features as vocally expressed languages.

2.2 Kinect

Kinect is voice and body recognition sensor that serves as controller for the XBOX gaming console, giving the users the ability to play games through voice and body gestures, without wearing or caring additionally accessories to track their body movements [11]. It features an RGB video camera, a depth sensor for 3D representation of the environment, a multi-array microphone for voice recognition and a property Microsoft software that enables human body recognition [1]. The software for skeletal tracking (tracking of the various body parts) is based on the data that comes from depth camera. It allows tracking of up to six people in its field of view, while offering full skeleton tracking for two of them. In full skeleton tracking mode, twenty (20) joints are tracked, while in seated mode, half of them [12]. The accuracy depends on the distance of the person from the camera where the distance from 1.2 meters to 3.5 meters gives the most accurate position of skeleton joints [12]. Joints are recognized as 3D points (with x, y, z coordinates), with the Kinect placed at the origin of the coordinate system and from Kinect viewpoint: positive z-axis increases towards the user, positive y-axis increases upward and positive x-axis increases to the left [13]. This is shown in figure 1.
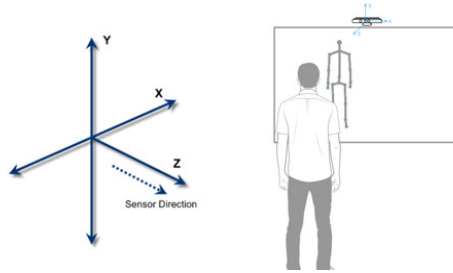
Figure 1 - Kinect skeleton coordinate space

The process of enabling skeleton tracking feature at Microsoft started by taking depth images for which body parts were known and used them to generate more depth images in order to cover all body types and positions [14]. So, 100,000 real depth images taken using motion capture system were used to generate one million depth images using computer graphic techniques [14]. Randomized decision forests were used to match these newly generated images with corresponding body parts [14]. Finally at the end these body parts were transformed into joint positions by using mean shift algorithm that finds densest region in each body part and considers it as joint position [14]. Hand shape is crucial in sign making and hand shape recognition is not offered through the standard SDKs. SigmaNIL is a framework that is capable of recognizing hands and has been used to enable the hand shape recognition. Its use will be detailed further in sections below that describe how the system was implemented.

2.3   Sign Language Recognition Process

In the literature there are many different approaches for the process of automatically recognizing sign language, but in general the process has three phases: the data acquisition phase, followed by the feature extraction phase and finally the analysis phase which concludes whether or not the combination of features that describe a particular sign happened. The process is depicted in figure 2.
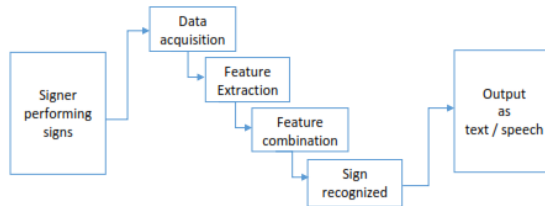
Figure 2 - Sign language recognition process

During the data acquisition phase, information about the movements of the user are collected. This may be achieved in one of two ways: the first, requires from the sign language performer to wear sensors (usually accelerometers) in body parts that are involved in sign making. These are usually gloves with embedded sensors that transmit the hand configuration to the connected computer wirelessly. These kind of

systems have been developed in [15], [16], [17] where gloves were connected to a computer for sign information transmission. Another system that required gloves was built in [18] but this time the glove transmitted information wirelessly to the computer. This approach is cumbersome and not convenient although it transmits accurate information. The second approach is more advanced as it is based on computer vision and offers a more natural interaction. Sign features in this category are derived by image and sensor data processing (from cameras). Such system has been developed in [19] where two ways were used for feature extraction: the first extracted features based on hand's skin color while the second required colored gloves for better segmentation of the hand. Computer vision techniques allows for more flexibility. For example the developed system [20] used two experiments: in the first the camera was placed on the desk and in the second the camera was placed on a cap worm that user had to wear; in another system [21] three camera were placed orthogonally to extract features in 3D. The second approach (vision based) is more favorable also in the cases were non-manual features of sign language are considered [22], [23] since the first approach would require wearing of cap worm for sensing head movements and facial expressions. The Kinect device fits in the second category. Kinect as presented earlier has two cameras: an ordinary one and a depth camera. Usually in SLR systems Kinect's depth camera is used to infer human skeleton positions. Kinect has been used in CopyCat game developed by CATS which tries to improve memory skills of HI children by allowing them to practice sentences in ASL [24]. In another system Kinect is part of humanoid robot NAO and is used to teach sign language to HI children: the robot performs signs, encourages the children to imitate and then watches whether the sign was performed correctly [25]. Some other SLR systems with Kinect are [26], [27], [28], [29], [30], [31] and [10]. The second phase (feature extraction) involves the use of the information stream that was acquired at the first phase, in order to identify important features (like the hand positions). The final phase uses the features that where extracted at the second phase in order to conclude which sign was performed. As it was discussed previously, sign language is multimodal, where information is conveyed simultaneously from different modes at the same time (hands move at the same time, may have different hand shapes, while at the same time head is moving and different facial expressions are present). To achieve correct analysis in the third phase usually Hidden Markov Models (HMM) are used [32]. HMM have been successfully used in speech recognition [33], but for sign recognition they have to be adopted for multimodal use, and this may be challenging [32]. Such systems have been developed in [16], [19], [20], [21], [22] and [10]. An alternative could be Artificial Neural Networks (ANNs) [32]. Sign combination can be seen as patterns. ANNs try to mimic the brain functionality for pattern recognition and simultaneous processing of information from different channel [34]. Some of the systems that use ANNs are [17], [35], and [28]. A more complete cover of SLR systems is available at [32], [35] and [36]. However this method has not had a great success. Our system does not use any of the above mentioned methods but rather employs an ad-hoc approach since the interest is inclusion and testing on hand shape feature in Sign Language Recognition (SLR) systems that are build using Kinect. This method has the advantage that it is simple in

getting results, but within a limited range of movements, as each hand gesture has to explicitly programmed at the system.

3 The Developed System

3.1 Description of the Arrangement

The idea is to facilitate the communication between HI and NH people by building a translation system that uses Kinect. The system will play the role of human sign language interpreter that understands the sign language and translates it into speech for NH people. Idea visual representation of the proposed arrangement is illustrated in figure 3. Three parties will be involved: the HI person, system that consist of Kinect and a computer and NH person. The HI person perform signs in front of the Kinect. Kinect tracks the HI person and transmits sign information to the computer. Computer analyzes the sign and provides its meaning as speech to NH person.
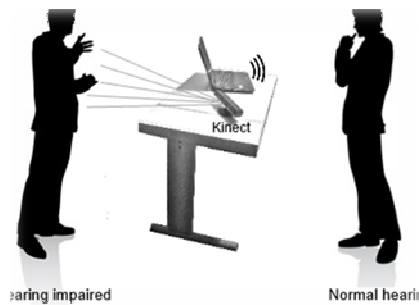
Figure 3 – Proposed system. The Hearing Impaired person performs signs in front of a Kinect, which translates the signs to text in a computer screen for a Normal Hearing person to read.

3.2 Constraints

As discussed in section 2.1, signs are composed from manual and non-manual features (such as eye blinking). In general, manual features involve using hands while non-manual features involve facial expressions, body postures and head movements. Due to time limitations the developed system deals only with recognition of manual features. Further more, the system is thought to be functional only in in-door environments, since the infrared lights that Kinect employs as part of depth measuring are destructed by sun-lights in outdoor environments.

3.3   System Design

From physical viewpoint the system consists of two components: the Kinect device and a computer. Kinect senses the scene and transmit the information to the computer where they analyzed and concluded which sign was performed. The developed software components are located on the computer only. The Kinect sensor supplies raw RGB and depth sensor's data which then are processed by software components residing on the computer. The software components are organized in a layered architecture. At the lowest layer are base SDKs that take raw data and construct RGB and depth images and then from depth images they infer skeleton positions. Going one layer up, the SigmaNIL framework utilizes these data to generate hand related features (e.g. the shape of the hand). Finally at the top of the hierarchy resides the developed components of the system and these components utilize skeleton positions and hand shape features to derive additional features like hand location and hand movement direction in order to describe a sign. In this layered architecture the communication is event driven. SigmaNIL waits for events from underlying SDKs and the developed application waits for events from SigmaNIL. The developed system consists of these components: the hand shape recognition component, hand position relative to other body parts, hand movement direction and text-to speech.

3.3.1   Hand Shape Recognition

Many alphabet letters and digit numbers of a sign language can be recognized if only the hand shape is recognized. The process of hand shape recognition passes by in two phases: first the selected hand shapes to be recognized are recorded, labeled and put in the SigmaNIL Training tool. This is a build-in tool that uses an algorithm for hand shape recognition and as a final output gives a database file with rules for distinguishing different hand shapes. Then this file is utilized in the software, where for each recognized shape a unique label is generated. SigmaNIL framework is in beta stage, with lack of documentation, support and furthermore it does not offer hand shape recognition in both hands (limited to one). Nevertheless, a way has been found to get recognition of hand shape in both hands at the same time. This was done by using the same database of hand shapes but introducing two additional SigmaNIL engines (segmentation and shape) for other hand, a solution similar (not identical) as for recognition of shapes with one hand and of course with a side effect on decreasing the performance, however not very noticeable. Hand shape recognition is provided by SigmaNIL framework. However the shapes that it will recognize must be pre-selected and they have to go through a training phase before they are able to be recognized (the training phase is elaborated more in system implementation section). Since recognition of all hand shapes is a long process a set of twenty (20) hand shapes were selected. Among selected shapes are those that enable recognition of basic sign language components and the shapes that may be used in different signs. However in a fully developed system, more hand shapes can be used. Using hand shape recognition the system can recognize all letters and numbers.

3.3.2 Hand Position in Relation to Each Other and to Other Body Parts

Hand shape identification alone, is not enough for recognition of larger sets of signs. Another important aspect is hand position relative to other body parts. This involves checking the position of hands relative to each other (are they near or are they touching and in which direction), to the head or to the chest and many more. SigmaNIL does not offer this feature, and therefore it was necessary to augment with appropriate code. Detection of hand position relative to each other was done by using a 2D representation of hands as they are presented in computer graphic coordinative system. A rectangle border around each hand was defined. The intersection of the two borders around the hands was used to define whether the hands touched each other. Distance between hands is measured as Euclidean distance between points of interest and if the distance decreased bellow a threshold it was concluded that the hands intersected. Figure 4 shows how the proximity of two hands is calculated: first two rectangular borders are drawn around each hand (figure 4.a and 4.b). The top two corners that are closer to each other are used then as points of interest, and the x and y difference is then calculated (figure 4.a and 4.b). To calculate the intersection between the two areas, the same borders are used. This is shown in figure 4.c and 4.d.
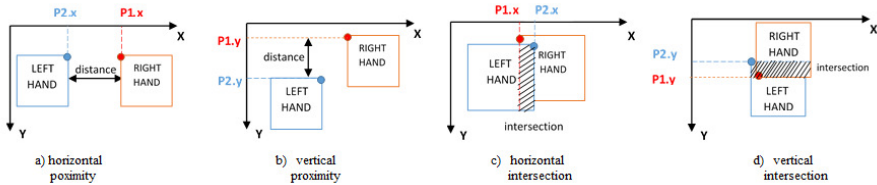
Figure 4 – Calculation of hands proximity: a) horizontal, b) vertical and hands intersection c) horizontal and d) vertical

In order to identify the relative position to other body parts (like the head, the chest and not between hands) the SigmaNIL framework had to be modified to draw borders around three body parts of special significance: the head, the torso and the inner torso (see figure 5.a). However these are larger areas compared to the hands and determining hand relative positions is done by introducing an additional method, by identifying that a hand is completely with in the border of a larger area. This is shown in figure 5.b.
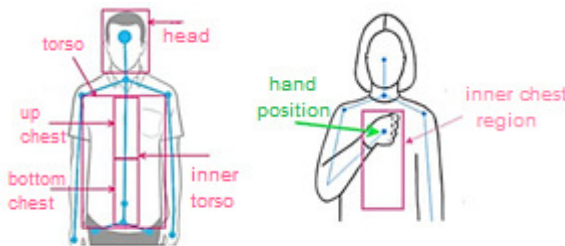
Figure 5 – Hand position relative to body parts: a) important regions and b) determining hand position

3.3.3 Hand Movement Direction

Hands can move in one of three directions in space: right-left, up-down, and away-towards the user. These are seen as basic movements of hands although more advanced movements exist. Movement can be understood as change of hand position from one location to the other. Since hand position is provided as skeleton position from Kinect SDK, implementation of this feature requires first understanding of Kinect skeletal coordinate system. The origin of the coordinative system is at the Kinect center as it is shown in figure 1. Each skeleton joint is represented by three dimensional (x, y, z) point. Furthermore as illustrated in figure 1, the x dimension represents the horizontal direction, the y dimension represents the vertical direction and the z dimension represents the direction from the Kinect towards the users' body. So movement of the hand along the x dimensions can be represented as *left* or *right* movement, movement along the y dimension as *up* or *down* movement and movement along the z dimension as *away from user* or t*owards the user* movement.

3.3.4 Alphabet, Number, Word and Sentence Recognition

The alphabet and number recognition can be demonstrated by showing the recognition of signs that involve one and two hands. For recognition of one hand signs, hand shape recognition is enough. This is showed in figure 6 that shows recognition of letter C (figure 6.a) and number one and number five (figure 6.b). For recognition of two hand signs within alphabet and number category, relative position of hands is also important. The figure 6 shows also the recognition of number nine (figure 6.c), letter A (figure 6.d) and letter G (figure 6.e).


a) letter C   b) nr. 1 and 5   c) number 9   d) letter A   e) letter G
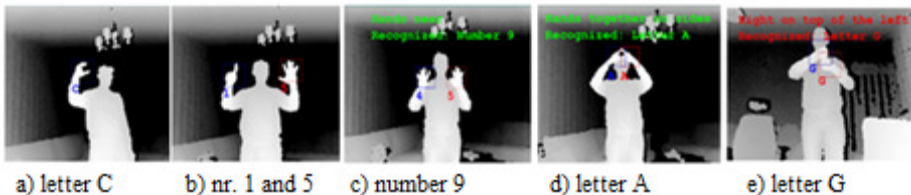
Figure 6 – Alphabet and number recognition: a) letter C, b) number 1 and number 5, c) number nine, d) letter A and e) letter G

The word recognition is demonstrated by showing the recognition of the word "HELLO". This word is a composition of a particular hand shape and a move in a particular direction starting from a particular body position. The hand shape, movement direction and hand location are shown in figure 7. To recognize this word, first is checked if the hand has this particular shape (similar to number four). Next is checked the location of the hand in relation to the head. If the sign is performed with the right hand as in figure 7, the hand must touch the right side of the head. Then after the hand touches the head, it must move in two directions at the same time: away from the user (toward Kinect) and a little up. The implementation is done by starting a timer when the hand touches the head. Then if within four seconds the hand moves

away from the user and little up it is concluded that sign for "HELLO" word is implemented. In a similar way other words were programmed. Composing sentences is then an issue of recognizing consecutive words. Sentences are treated as sequential occurrence of words that happen within a period of time. Sentence recognition is demonstrated through following example. The figure 8 shows sequential occurrences over the time of the signs that represent two words. The first is the sign for "HELLO" word and the second is the sign for "DAUGHTER" word, both in KLS. When they happen sequentially within a period of time they form the sentence "HELLO DAUGHTER"The recognition of sentence "HELLO DAUGHTER" is done by starting the timer when the HELLO signs is performed and if within five seconds if sign DAUGHTER happens, it is concluded that sentence "HELLO DAUGHTER" was performed. The last component provided is sign to speech translation. Each recognized sign is translated into spoken language and outputted as a text and as speech. This section concludes the functionalities provided in this system.
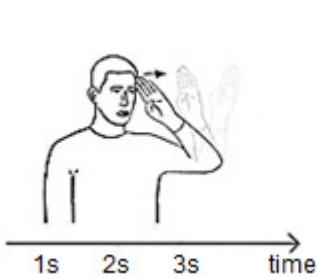
Figure 7 - Recognition of *Hello* word
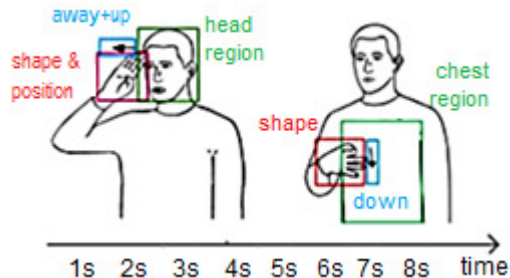
Figure 8 - Recognition of sentence: *Hello Daughter*

3.4   Testing and Evaluation of the System

The functionalities for which the system is tested are: number, alphabet letters, words and sentence recognition and text to speech as well. The system was tested under different light conditions, in real time. The system was tested from two (2) native speakers of sing language and one (1) non-native speakers of sign language. The testers were of different ages, body types and sizes. Different testers assure testing for different sizes and shapes of hands thus providing more insights into using the hand shape feature in systems for sign language recognition. The system was tested for nine numbers, fifteen alphabet letters, four words and one sentence. Testers performed each signs ten (10) times. The non-native speakers had to be trained to perform the signs whereas non-native speakers were required just to perform the signs. The results from testing are shown in table 1 and graphically presented in figure 9. In table 1 the second column lists the tested signs and three other columns lists the accuracy results from three different system testers. Each time the tester performed the sign, it was observed whether the system correctly recognized it and if it did it was counted. As it is shown in table 1, the sign for number one was performed ten times by each tester and in the case of the first tester the system recognized correctly it each time, in case

of second tester the system recognized correctly it nine (9) times and in the case of the last tester the sign was recognized correctly eight (8) times. In the case of the sign for letter A and B the tests were canceled for native speakers because they were not recognized at all after initial testing with non-native tester. This happened because of hand occlusion where one hand was occluded by the other and not observable by Kinect viewpoint.

Table 1 - Results from testing

Sign	Non-native tester	Native tester 1	Native tester 2
Number 1	10/10	9/10	8
Number 2	10/10	7/10	9/10
Number 3	10/10	10/10	10/10
Number 4	9/10	8/10	8/10
Number 5	10/10	10/10	10/10
Number 6	10/10	1/10	3/10
Number 7	10/10	7/10	9/10
Number 8	10/10	9/10	10/10
Number 9	8/10	10/10	10/10

Sign	Non-native tester	Native tester 1	Native tester 2
Letter A	5/10		
Letter B	5/10		
Letter C	10/10	9/10	10/10
Letter E	8/10	8/10	10/10
Letter L	10/10	10/10	10/10
Letter LL	8/10	10/10	10/10
Letter O	10/10	10/10	10/10
Letter U	9/10	7/10	4/10

Sign	Non-native tester	Native tester 1	Native tester 2
Word: OK	8/10	6/10	6/10
Word: Hello	9/10	10/10	10/10
Word: Daughter	8/10	8/10	8/10
Word: Are you OK ?	7/10	10/10	8/10
Sentence: Hello daughter	4/10	8/10	10/10

The graph in figure 9 presents the recognition accuracy for each component of sign language for each tester. It can be seen that accuracy is higher in case of non-native speaker for number and alphabet letters while it decreases in case of words and sentences.
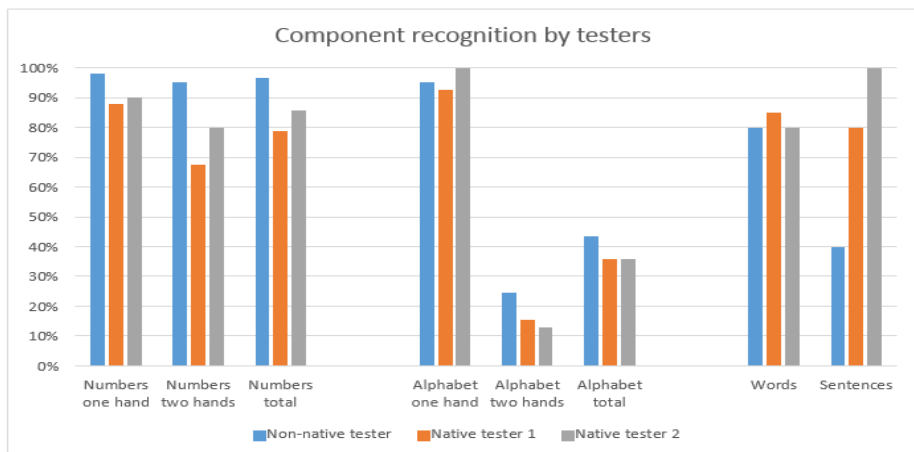
Figure 9 – Graphical representation of test results for each component of sign language for each tester

Higher accuracy for non-native speaker can be explained from the fact that non-native speaker was trained to perform signs from system developers and thus un-intentionally influenced the way signs were performed. In the other side, native speakers have interpersonal variation in performing same signs and thus the accuracy varied. However in the case of more advanced signs like words or sentences the system recognition accuracy with native speakers was higher. One possible explanation is that more complicated signs are harder to learn for non-native speakers.

As mentioned earlier in this section the system had difficulties in correctly recognizing signs that involved both hands and especially cases where hands occluded each other. This is shown also in the graph where the recognition accuracy for signs with one hand is much higher.

3.4.1 Evaluation of the System

The system is able to recognize numbers from 1 to 9 in sign language with high accuracy. However not all numbers of sign language were recognized such as numbers starting from ten and up. In addition, the KSL alphabet consist of thirty-six alphabet letters. From them fifteen were implemented and tested. The system showed high recognition accuracy for alphabet letters that involved one hand. In cases where two hands are involved the recognition accuracy drops. This holds for signs when hands touch each other, while in the signs where hands do not touch (near only) the accuracy is similar to alphabet letters made with one hand. Furthermore the alphabet letters that were combination of other alphabet letters were not implemented at all. The system implements a limited vocabulary of words and sentences. More specifically four words and one sentence were implemented and tested. Although recognition accuracy is higher and at acceptable levels, the system does not employ common methods (HMMs or ANNs) for word and sentence recognition that could have increased recognition accuracy. As an additional consequence the system is not easily scalable for addition of new words and sentences. Sign languages incorporate non-manual features (facial expression, body posture, hand movement) that were not taken into consideration in this project. The capability for continuous recognition of sign language communication is not provided. Continued communication in sign language does not require recognition of alphabet letters, since sequential combination of words is what forms sentences that are used for continuous communication. Sentence recognition was tested for one sentence only and followed a simple algorithm that is not scalable for inclusion of new sentences. Although SigmaNIL proved to have good potential for hand shape recognition it had to be modified in order to provide hand shape recognition in both hands and also to proivde tracking of other joints than hands. These two modification resulted in decrease on system performance (fortunatley not noticable by users) and deviation from coding standards (implicatoin on future system releases). Beside these facts it is very promising framework to be used in building SLR systems and hopefully it will be enhanced in next releases.

4   Future Work

The system can be improved further either by enhancing existing futures or by providing new ones. The recognition accuracy of hand shape component can be increased further by playing with parameters of the algorithm that is available through SigmaNIL training tool. Furthermore recognition accuracy can be increased by employing proven methods like HMMs, ANNs and other that were used in similar systems (refer to section 2.3). Vocabulary of recognized hands shapes can be

increased to include large data set. The hand position relative to body feature can be enhanced further by defining all necessary hand position to be detected and then by enhancing the algorithms for detection of these positions. The 2D and 3D feature were used to detect different hand positions. Implementing algorithms that are based only in 3D futures would increase the performance by eliminating the need of conversation of 3D into 2D and then using 2D for comparison purposes. Hand movement feature also can be enhanced further. It provides recognition of six basic hand movements that can be combined to detect more movement directions. Identification of other movement directions involved in sign making and their detection would provide the possibility for recognition of larger vocabulary of signs. While numbers, alphabet letters and words recognition can be enhanced by enhancing these features, sentence recognition requires implementation of more sophisticated mechanism like HMMs, ANNs and other methods, in order to provide continuous SLR. Finally the system can be enhanced further by providing new features, like bi-directional translation and tracking of non-manual features. While incorporation of non-manual features is important, providing the speech to signs translation will definitively provide better communication between hearing impaired and normal hearing people.

5   Overview and Conclusions

The Kinect sensor was investigated for building a system that translates sign language to spoken language. While sign language uses manual and non-manual signs to conveying meaning in this work only manual signs were investigated. Manual signs are composition of hand shape, hand location, hand orientation, and hand movement. The developed system incorporates these components and uses them to recognize and translate to speech: numbers, alphabet letters, words and sentences from Kosovo Sign Language. Four components of this language were recognized and tested: numbers from 1 to 9, 15 alphabet letters, four words and one sentence. The system is tested by two (2) native speakers of sign language and one (1) person without prior knowledge on SL to observe the effect of sign language proficiency in recognition accuracy of the system. Results of testing system with native and non-native speaker showed that proficiency with sign language and interpersonal variation are important factors that have impact on recognition accuracy. Thus recognition accuracy of non-native speaker was higher for alphabet letters and numbers explained by the fact that they were easy to learn and since the tester was trained by system developers, they possibly influenced the tester in performing signs in the way system expects. In the other side recognition accuracy of native speakers for words and sentences was higher and can be attributed to their proficiency in sign language while variation in the accuracy to interpersonal variations. In general the system achieved higher recognition accuracy for one hand signs and lower recognition accuracy for signs where both hands were involved. While the recognition accuracy for static signs was high (around 90%), further work is needed in order to improve recognition accuracy of hand shape component when it is combined with hand movement and hand location features. In the other side, although Kinect is well suited for feature

extraction, proper software recognition method has to be followed in order to understand continuous communication in sign language. Finally as most of conducted researches do not incorporate the hand shape component, the incorporation and investigation of it in this work is seen as main contribution in the area of Sign Language Recognition systems with Kinect.

References

[1] Microsoft, "Kinect for Windows, Sensor Components and Specifications," 2013. [Online]. Available: http://msdn.microsoft.com/en-us/library/jj131033.aspx. [Accessed: 06-Dec-2013].

[2] M. Andersen, T. Jensen, and P. Lisouski, "Kinect depth sensor evaluation for computer vision applications," 2012.

[3] SigmaNIL, "SigmaNIL: Technical documentation of the framework." [Online]. Available: http://www.sigmanil.com/docs/SIGMANIL API DOCvs1.htm. [Accessed: 06-Dec-2013].

[4] W. C. Stokoe, "Sign language structure: an outline of the visual communication systems of the American deaf. 1960.," *J. Deaf Stud. Deaf Educ.*, vol. 10, no. 1, pp. 3–37, Jan. 2005.

[5] K. Emmorey, *Language, cognition, and the brain: Insights from sign language research*. Psychology Press, 2001.

[6] M. P. Lewis, G. F. Simons, and C. D. Fening, "Ethnologue: Languages of the world (17th edition)," SIL international Dallas, TX, 2013.

[7] S. K. Liddell, *Grammar, gesture, and meaning in American Sign Language*. Cambridge University Press, 2003.

[8] R. Battison, "Lexical Borrowing in American Sign Language.," 1978.

[9] U. Bellugi and S. Fischer, "A comparison of sign language and spoken language," *Cognition*, 1972.

[10] S. Lang, M. Block, and R. Rojas, "Sign language recognition using kinect," *Artif. Intell. Soft Comput.*, pp. 394–402, 2012.

[11] Microsoft, "Kinect Fact Sheet," 2010. [Online]. Available: www.microsoft.com/enus/news/presskits/xbox/docs/kinectfs.docx.

[12] Microsoft, "Kinect for Windows, Human Interface Guidlines." [Online]. Available: http://msdn.microsoft.com/en-us/library/jj663791.aspx. [Accessed: 06-Dec-2013].

[13] Microsoft, "Kinect Coordinate Space." [Online]. Available: http://msdn.microsoft.com/enus/library/hh973078.aspx#Depth_Ranges. [Accessed: 06-Dec-2013].

[14] J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore, "Real-time human pose recognition in parts from single depth images," *Commun. ACM*, vol. 56, no. 1, pp. 116–124, 2013.

[15] T. Takahashi and F. Kishino, "Hand gesture coding based on experiments using a hand gesture interface device," *ACM SIGCHI Bull.*, vol. 23, no. 2, pp. 67–74, 1991.

[16] C. Lee and Y. Xu, "Online, interactive learning of gestures for human/robot interfaces," in *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on*, 1996, vol. 4, pp. 2982–2987.

[17] J. Weissmann and R. Salomon, "Gesture recognition for virtual reality applications using data gloves and neural networks," in *Neural Networks, 1999. IJCNN'99. International Joint Conference on*, 1999, vol. 3, pp. 2043–2046.

[18] K. Thomas, "Glove lends the deaf a hand," *USA Today. Retrieved Oct.*, vol. 10, p. 2002

[19] T. Starner and A. Pentland, "Real-time american sign language recognition from video using hidden markov models," in *Motion-Based Recognition*, Springer, 1997

[20] T. Starner, J. Weaver, and A. Pentland, "Real-time american sign language recognition using desk and wearable computer based video," *Pattern Anal. ...*, vol. 20, no. 12, pp. 1371–1375, 1998.

[21] C. Vogler and D. Metaxas, "ASL recognition based on a coupling between HMMs and 3D motion analysis," in *Computer Vision, 1998. Sixth International Conference on*, 1998, pp. 363–369.

[22] D. Kelly, J. Reilly Delannoy, J. Mc Donald, and C. Markham, "A framework for continuous multimodal sign language recognition," in *Proceedings of the 2009 international conference on Multimodal interfaces*, 2009, pp. 351–358.

[23] P. Dreuw, H. Ney, G. Martinez, O. Crasborn, J. Piater, J. M. Moya, and M. Wheatley, "The SignSpeak Project - Bridging the Gap Between Signers and Speakers," 2010.

[24] Z. Zafrulla, H. Brashear, and H. Hamilton, "American sign language recognition with the kinect," *Work*, pp. 279–286, 2011.

[25] H. Kose, R. Yorganci, E. H. Algan, and D. S. Syrdal, "Evaluation of the Robot Assisted Sign Language Tutoring Using Video-Based Studies," *Int. J. Soc. Robot.*, vol. 4, no. 3, pp. 273–283, Mar. 2012.

[26] F. Huang and S. Huang, "Interpreting American Sign Language with Kinect," pp. 1–5, 2011.

[27] E. Rakun, M. F. Rachmadi, K. Danniswara, and others, "Spectral domain cross correlation function and generalized Learning Vector Quantization for recognizing and classifying Indonesian Sign Language," in *Advanced Computer Science and Information Systems (ICACSIS), 2012 International Conference on*, 2012, pp. 213–218.

[28] F. Trujillo-Romero and S.-O. Caballero-Morales, "3D data sensing for hand pose recognition," in *Electronics, Communications and Computing (CONIELECOMP), 2013 International Conference on*, 2013, pp. 109–113.

[29] M. Oszust - "Polish sign language words recognition with Kinect," in *Human System Interaction (HSI), 2013 The 6th International Conference on*, 2013, pp. 219–226.

[30] H. Takimoto e "A Robust Gesture Recognition Using Depth Data."

[31] D. Capilla, "Sign Language Translator using Microsoft Kinect XBOX 360 TM," *Dep. Electr. Eng. Comput. ...*, 2012.

[32] H. Cooper, B. Holt, and R. Bowden, "Sign language recognition," in *Visual Analysis of Humans*, Springer, 2011, pp. 539–562.

[33] P. Dreuw and D. Rybach, "Speech recognition techniques for a sign language recognition system," …, Figures 7 and 8 have some problems.2007.

[34] Y. F. Admasu and K. Raimond, "Ethiopian sign language recognition using Artificial Neural Network," in *Intelligent Systems Design and Applications (ISDA), 2010 10th International Conference on*, 2010, pp. 995–1000.

[35] P. Vijay and N. Suhas, "Recent Developments in Sign Language Recognition: A Review," *irdindia.in*, no. 2, pp. 21–26, 2012.

[36] M. Majid and J. Zain, "A REVIEW ON THE DEVELOPMENT OF INDONESIAN SIGN LANGUAGE RECOGNITION SYSTEM," *J. Comput. Sci.*, vol. 9, no. 11, pp. 1496–1505, 2013.

A Multilayer Comparative Study of
XG-PON and 10G-EPON Standards

Charalampos Konstadinidis[1], Panagiotis Sarigiannidis[2], Periklis Chatzimisios[1], Paschalis Raptis[1], Thomas D. Lagkas[3]

[1]Department of Informatics, Alexander TEI, Thessaloniki, Greece
e-mail: {chrkon, peris, praptis}@it.teithe.gr
[2]Department of Informatics and Telecommunications Engineering, University of Western Macedonia, Kozani, Greece
e-mail: psarigiannidis@uowm.gr
[3]Computer Science Department, The University of Sheffield Internationally Faculty, CITY College, Thessaloniki, Greece
e-mail: tlagkas@city.academic.gr

Abstract. The purpose of this paper is to provide a multilayer review of the two major standards in next generation Passive Optical Networks (PONs) and technologies, the ITU-T 10-Gigabit-capable PON (XG-PON) and the IEEE 10 Gigabit Ethernet PON (10G-EPON). A study and a discussion on the standards are performed. The main intention of this paper is to compare XG-PON and 10G-EPON, mainly in terms of physical and data link layers. The paper answers the question of what are the common elements and the basic execution differences of the two standards. Moreover, critical points are raised regarding the Dynamic Bandwidth Allocation (DBA) schemes of both standards. Special focus is also pointed in the coexistence of XG-PON and 10G-EPON. Finally, the paper includes a discussion of open issues and continuing research regarding the two standards.

Keywords: Passive Optical Networks (PONs), optical access network, Fiber-To-The Home (FTTH), XG-PON, 10G-EPON

1 Introduction

Optical networks penetrate access networks in an increasing rate attributed by the extensive growth of Internet traffic as well as the increased use of bandwidth intensive applications. Cutting-edge multimedia services, such as Ultra High Definition (HD) video, are leading to even higher bandwidth requirements. Compared to copper, optical fiber can provide higher bandwidth over a longer distance. In order to partially fulfill the aforementioned requirement, telecommunication companies have been already deploying Fiber-To-The-x (FTTx) networks in various parts of the world. The concept of the different variations of the FTTx technology, namely FTTH (Home), FTTC (Curb) and FTTB (Building) is the provision of fast fiber connections in high proximity to the end user's premises. FTTx networks primarily constitute Passive Optical Networks (PONs) that are considered as one of the most promising
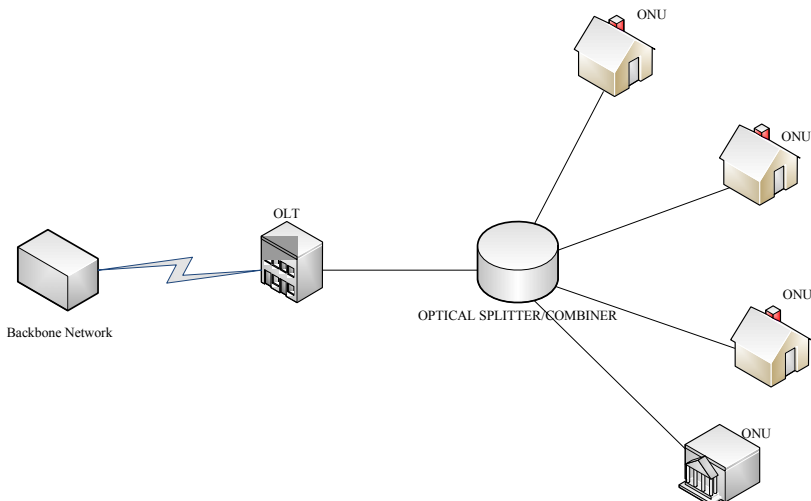
Fig. 1. The generic tree topology of a PON.

alternatives to dominate broadband access, due to their cost-effective and practically limitless bandwidth potentials. PONs are widely deployed in FTTx technologies, creating optical light-paths without incorporating optical-to-electrical conversions [1]. A generic view of a PON is depicted in Fig. 1.

ITU-T Gigabit-capable PON (GPON) and IEEE PON (EPON) are the two competing systems since 2009 having both of them the advantage to offer more bandwidth per subscriber than their predecessor Broadband PON (BPON). EPON has been proved a success mainly in countries in Eastern Asia, such as China, Korea and Japan, while GPON has been widely deployed in the North America.
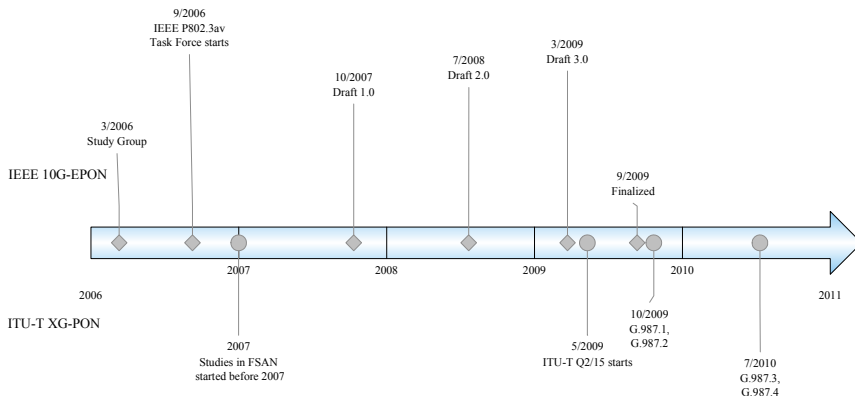
Fig. 2. Standardization timeline for 10 Gigabit PONs in IEEE and ITU-T [3]

1.1 10 Gigabit PONs

PON systems of 10 Gbit/s have been recently introduced to cover future high bandwidth demands and well as residential and backbone services [2,3]. The latest related IEEE and ITU-T standards are backwards compatible with previous PON generations, allowing upgrading in a progressive manner with minimal financial investment on the Optical Distribution Network (ODN) and minimal effect on existing users' operation. The standardization of IEEE 10G-EPON began in 2006 and ended in late 2009. On the other hand, ITU-T XG-PON has received its final form in 2010. The timeline for the standardization of IEEE 10G-EPON and ITU-T XG-PON is shown in Fig. 2.

   IEEE and ITU-T along with the Full Services Access Network (FSAN) group have recently defined their 10 Gbit/s solutions, namely IEEE 802.3av, 10GE-PON, and ITU-T XG-PON, respectively, in an effort to support the upcoming bandwidth increase over existing ODNs. The P802.3av Task Force has been standardizing 10 G-EPON, developing the IEEE 802.3av 10GE-PON which provides symmetric (10 Gbit/s downlink and uplink) and asymmetric (10 Gbit/s downlink and 1 Gbit/s uplink) connections. The asymmetric traffic generated by IP video services is expected to be supported by the latter operation mode.

   Regarding the related work of the FSAN group, its efforts are focused on examining next generation approaches that facilitate high provision, increased split ratio, and extensive reach. Under the consideration of Next Generation Access (NGA), which is mainly about technologies compatible with existing GPONs, XG-PON1 has been developed by ITU-T and FSAN; the supported asymmetrical downlink/uplink bandwidth capacity is 10/2.5 Gbit/s. The definition of XG-PON1 can be found in the ITU-T G.987.x series of recommendations. Some key characteristics of the XG-PON1 networks include security enhancements through authentication of management messages and power conservation through switching down parts or all of the Optical Network Units (ONUs) [4].

2 XG-PON Overview

After standardizing PON networks operating at 1 Gbit/s in ITU-T G.984.x Recommendation series, efforts have been made to standardize the 10 Gbit/s capable PONs (were finalized in October 2009 and published by ITU-T in March 2010). The ITU-T G.987.x series addresses the general requirements of 10 Gigabit capable passive optical networks in a way that promotes backward compatibility with the existing ODN that complies with the GPON systems. Furthermore, the XG-PON system utilizes Wavelength Division Multiplexing (WDM) defined in ITU-T G.984.x series and provides a seamless migration from Gigabit PON to XG-PON.

2.1 Physical Layer

The physical layer, also referred to as the Physical Media Dependent (PMD) layer is described in the ITU-T G.987.2 Recommendation and specifies a flexible optical fiber

access network capable of supporting high bandwidth requirements. A general requirement of XG-PON is to provide higher data rates than GPON combined with minimized costs. Therefore, backward compatibility with legacy GPON deployments was a major topic in the physical layer specifications. To achieve the backward compatibility and co-existence of GPON and XG-PON systems, the optical wavelengths that selected for XG-PONs were the "O-band" (for the upstream ranging from 1260 to 1280 nm) and the "1577nm" (for the downstream ranging from 1575 to 1580 nm).

XG-PON systems are divided into XG-PON1 (featuring a 2.5 Gbit/s upstream path) and XG-PON2 (featuring a 10 Gbit/s one). ITU-T G.987.x Recommendation series [5,6,7] only addresses XG-PON1. It seems that a natural progression from GPON to XG-PON1 and XG-PON2 is to come when the technology becomes more mature.

Having in mind that XG-PON systems should share the same optical distribution network as GPONs, the ODN characteristics are quite clear (28 dB maximum loss in the windows from 1260 to 1360 and from 1480 to 1580 nm). The FSAN operators identified two loss budgets for the XG-PON systems. The Nominal 1 class, with a 29 dB maximum loss and the Nominal 2 class with a 31 db maximum loss. The first one allows XG-PON to coexist with standardized GPON systems, while the second one supports coexistence with the super-standard 29.5 dB GPON systems.

2.2 Transmission Convergence (TC) Layer

The XG-PON data link layer (usually referred as Transmission Convergence - TC layer) is composed of three distinct sublayers: the XGTC framing sublayer, the XGTC PHY adaption sublayer and the XGTC service adaption sublayer [8]. The TC layer's main function is to provide transport multiplexing between the Optical Line Terminal (OLT) and the ONU, followed by other functions like adaption of client layer signal protocols, Physical Layer Operations And Maintenance (PLOAM), interface for Dynamic Bandwidth Allocation (DBA), and ONU ranging and registration. XGTC functions are realized through Transmission CONTainers (T-CONTs) [9], each of them identified by a unique Allocation ID (Alloc-ID) assigned by the OLT.

The XGTC service adaption sublayer is responsible for taking the user's payload and formatting them for transmission over optical network. In an XG-PON system, the Service Data Units (SDUs), which include the user data frames and high-level PON management frames, are transmitted in XGTC payload section using the XG-PON Encapsulation Method (XGEM system). The XGEM supports SDU fragmentation, encapsulation and delineation both in the downstream and upstream directions, and marks the individual flows of traffic (ports) so that they can be accepted by the appropriate client on the other side of the PON.

The XGTC framing sublayer is responsible for the construction and parsing of the overhead fields in both the transmitter and the receiver side. On the transmitter side, the framing sublayer accepts the XGEM frames from the XGTC service adaption sublayer and then constructs the downstream XGTC frame or the upstream XGTC burst by providing embedded OAM and PLOAM messaging channel overhead fields. On the receiver side, the framing sublayer accepts the XGTC frames or XGTC bursts, parses the overhead fields by extracting the incoming OAM information and PLOAM
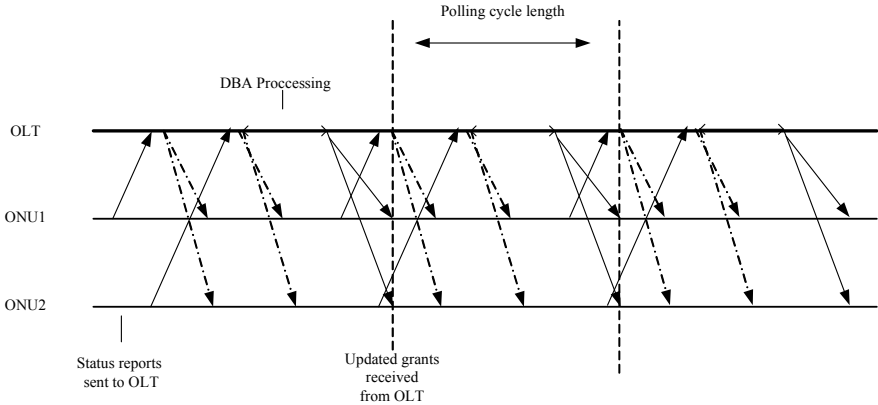
Fig. 3. The DBA process in XG-PON Systems.

messaging flows and then delivers the XGTC payloads to the service adaption sublayer.

The PHY adaption sublayer takes care of the low level coding of the TC frame over the physical channel trying to improve the detection, reception and delineation properties of the signal transmitted over the medium. Much of the work of the PHY adaption layer concerns the use of Forward Error Correction (FEC), a required feature for downstream and upstream directions. The use of FEC improves the effective sensitivity and overload characteristics of the optical receiver by introducing redundancy in the transmitted bit stream and allowing the receiver to operate under higher BER (Bit Error Rate) scenarios.

XGTC Encapsulation Method

The transmission of the SDUs (that include the user data frames and high-level PON management frames in the XTC payload section of the downstream XGTC frames and upstream XGTC bursts) is accomplished through the XGEM encapsulation method. The XGTC payload section is carried in the downstream frames or upstream bursts and contains one or more XGEM frames. Each XGEM frame contains a fixed size header carrying information such as Payload Length Indication (PLI), XGEM port-ID as well as last fragment indication and variable size XGEM payload field controlled by the PLI.

Dynamic Bandwidth Allocation in XG-PON

In a XG-PON system the OLT provides media access control for the upstream traffic. The basic idea is that each downstream frame comes with a BandWidth map (BWmap) that indicates the location for upstream transmissions by various ONUs in the corresponding upstream PHY frame. The header of downstream XGTC frame contains the BWmap field, which specifies a bandwidth allocation to a particular allocation ID (Alloc-ID), while in the upstream burst the allocation overhead is

composed of the Dynamic Bandwidth Report upstream (DBRu) structure and contains the Buffer Occupancy (BuffOcc) field that reports the total amount of SDU traffic.

XG-PON uses Point-to-Multi-Point (P2MP) connections between the OLT and the ONU. Fig. 3 illustrates the logical diagram of the DBA process. Due to the high available bandwidth in XG-PON, bandwidth allocation is based on the Service Level Agreements (SLA) where Quality of Service (QoS) can be granted according to the demand. Bandwidth is allocated per Transmission CONTainer (T-CONT) [9,10], which is the basic control unit for bandwidth allocation. Each T-CONT is indexed by an Alloc-ID. T-CONTs represent a logical communication link between the OLT and the ONUs, with every single ONU being able to assigned one or more T-CONTs. There are three different types of T-CONT feasible for dynamic bandwidth allocation, T-CONT types 2, 3 and 4. T-CONT type 2 is for on-off type traffic with well defined rate bound and strict delay requirements provisioned with assured bandwidth. This bandwidth has to be granted to the T-CONTs' traffic, if requested. If not used, bandwidth can be allocated to other T-CONTs, providing that it is available as soon as T-CONT type 2 requires it. T-CONT type 3 is provisioned with assured bandwidth and it can also be granted non-assured bandwidth if the entire assured bandwidth is utilized. It is suitable for variable rate, bursty traffic with requirements for average rate guarantee. T-CONT type 4 has no bandwidth guarantee but it has eligibility in best effort bandwidth sharing.

In XG-PON the DBA follows a strict priority hierarchy among the forms of assigned bandwidth:

    1)   Fixed bandwidth (highest priority)

    2)   Assured bandwidth

    3)   Non-assured bandwidth

    4)   Best-effort bandwidth (lowest priority)

Firstly, the OLT assigns the upstream bandwidth to the fixed bandwidth of each Alloc-ID. Secondly, the OLT allocates the assured bandwidth of each Alloc-ID as long as the Alloc-ID has enough traffic to consume the assured one. The OLT then satisfies the requirements of non-assured bandwidth to the eligible unsaturated Alloc-IDs until either all of them reach their saturation level, or the surplus bandwidth pool is exhausted. Finally, the OLT allocates the remaining bandwidth to the best-effort bandwidth components.

Because XG-PON adopts this strict hierarchy in bandwidth allocation, the received QoS of a request is determined by the T-CONT type that maps this request. The mapping must consider not only the QoS requirements of the application but also the traffic characteristics. For example, applications such as HDTV streams or video conferencing for business subscribers (who are willing to spend more to acquire a guaranteed QoS) are potentially mapped into T-CONT type 1.

3 The 10G-EPON Overview

The effective 1 Gbit/s symmetric data rate supported by the IEEE 802.3-2005 compliant EPON systems has been considered sufficient for a relatively short period

of time. However, the increasing demand of raw bandwidth and high capacity, resulted in the development of the 10G-EPON. The 802.3av PON standard [11] was developed to increase the data rate of EPON systems from 1 Gbit/s to 10Gbit/s and being compatible with the 10 Gbit/s Ethernet interface. There are multiple protocols for both 10G-EPON and EPON. As an addition to the Ethernet family of IEEE 802.3, EPON and 10G-EPON layering is very similar to that of Point-to-Point (P2P) Ethernet. The physical layer is connected to the data link layer using the Media-Independent Interface (MII) or the Gigabit Media-Independent Interface (GMII).

3.1 Physical Layer

The physical layer specifies the physical characteristics of the optical transceivers. Ethernet has the tradition of adopting mature low-cost designs to promote mass deployment. This philosophy has been the key to the tremendous commercial success of Ethernet. The Physical layer is subdivided into six blocks [12]:

1) MDI specifies the characteristics of the electrical signals which are received from or transmitted to the underlying medium
2) PMD specifies the basic mechanisms for exchange of data streams between the medium and PCS sublayer. The bottom part of PMD contains physical devices, like receiver and transmitter.
3) PMA sublayer specifies functions responsible for transmission, reception, clock recovery, and phase alignment.
4) PCS defines a set of functions which are responsible for converting a data stream received from GMII into codewords, which can then be passed through PMA and PMD and finally transmitted into the medium.
5) GMII specifies a standardized interface between the MAC and PHY layers. This is one of the major interfaces in the 802.3 stack allowing for modular interconnections of various PHY layers to MAC
6) RS maps MAC service primitives into GMII signals, effectively transfers data into PHY and vice versa. In the EPON architecture, RS plays also one more critical role: it is responsible for LLID insertion and filtering all data passing from MAC or PHY.

According to the specifications of 10G-EPON, it offers symmetric 10 Gbit/s at the downlink and the uplink, as well as asymmetric 10 Gbit/s downlink and 1 Gbit/s uplink data rates. Moreover, in 10G-EPON the OLT is equipped with dual rate receivers for 1G or 10G ONUs to be backwards compatible with the existing and widely deployed 1G –EPON. Additionally, the downlink transmission channels are divided for sending data and control information to 1G and 10G ONUs.

Allowing concurrent operation of 1 Gbit/s and 10 Gbit/s EPON systems was a major priority for the 802.3av standard. On the downlink, the 1 Gbit/s and 10 Gbit/s channels are divided based on wavelengths, with the 1 Gbit/s transmission using the 1480 to 1500 nm band and 10 Gbit/s spreads from 1575 to 1580 nm. On the uplink, there is an overlap; 1 Gbit/s uses the 1260 to 1360 nm band, whereas 10 Gbit/s spread from 1260 to 1280 nm.
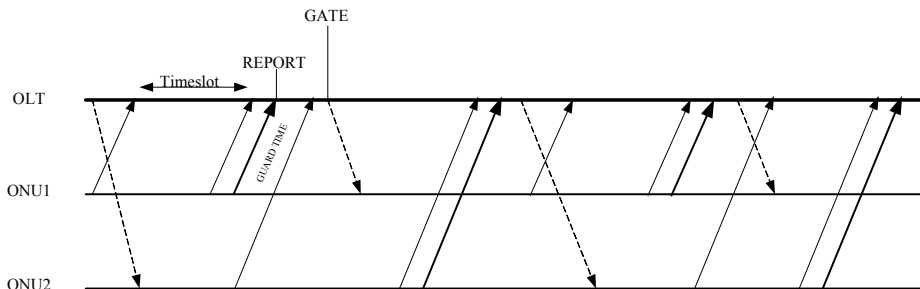
Fig. 4. The DBA process in 10G-EPON Systems.

3.2 Data Link Layer

The Ethernet layering architecture has differences between P2P and P2MP models. In the data link layer of 10G-EPON, in P2P Ethernet a mandatory Multipoint Media Access Control (MPMC) layer replaces the optional MAC sublayer. The MultiPoint Control Protocol (MPCP) is part of the MPCP layer and is employed to manage access of the 10G-EPON ONUs to the shared PON medium. It is noted that despite the fact that the OLT and ONU stacks are quite similar, MPCP in an OLT has the role of the master, while the MPCP entity in ONU functions as the slave.

MPCP is the protocol adopted to arbitrate the uplink transmission among the ONUs, defined by the IEEE 802.3ah task force. It does not involve a specific DBA scheme, but it allows the implementation of DBA schemes by facilitating information exchange necessary by the OLT to assign bandwidth to each ONU. MPCP controls the access to the P2MP topology architecture via message, status and timer. Its main functions are ONU bandwidth allocation, polling ONU bandwidth requests, reporting congestion to up-level, ONU auto-discovery, registration, and ranging[13,14]. It includes two 64-byte MAC control messages, GATE and REPORT [14].

Dynamic Bandwidth Allocation in 10G-EPON

There were no significant changes in the DBA scheme supported by GEPON systems suffered when transiting towards 10 Gbit/s EPONs. Due to the case of co-existence, the emerging 10 Gbit/s EPONs have the DBA operation based on the underlying MPCP sublayer. Thus, the DBA entity is responsible for scheduling two mutually cross dependent EPON systems, which use a common single upstream channel. In downlink, the DBA agent schedules the transmission of the GATE MPCPDUs independently, since WDM multiplexing is used to separate the 1 Gbit/s and 10 Gbit/s paths [13-16]. P2P emulation is achieved by a mechanism, which allows the medium to behave as a collection of P2P links. The emulation depends on tagging Ethernet frames with a unique identifier for each ONU called the Logical Link ID (LLID), which is embedded in the frame preamble. Fig. 4. shows the 10G-EPON DBA process.

With DBA, the OLT assigns the bandwidth to different ONUs based on the data they have to send rather than a static allocation per ONU. The 10G-EPON uses

REPORT messages from the ONUs to inform the OLT of their current bandwidth
needs. Their bandwidth needs are reported in terms of the number of characters they
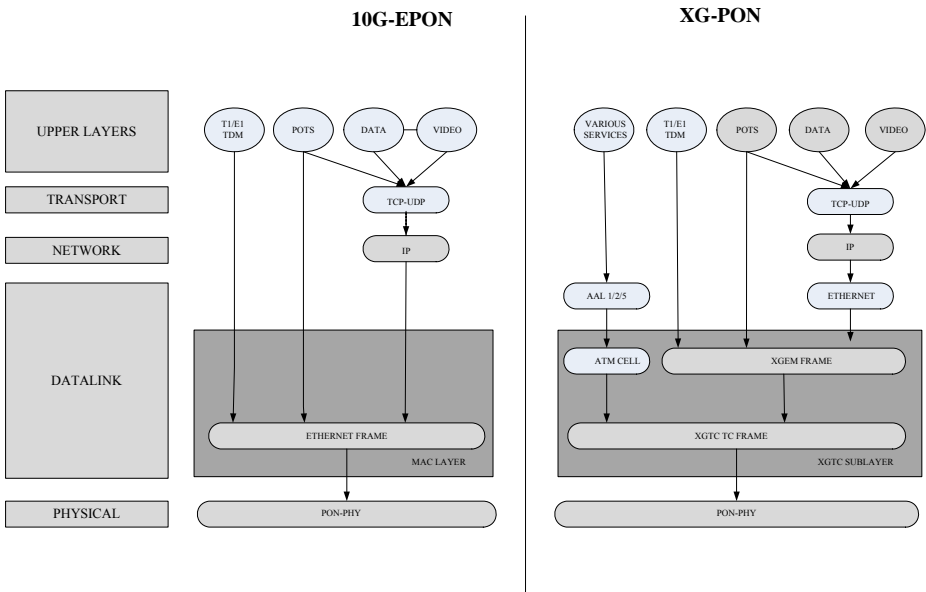
Fig. 5. The multilayer comparison between the two standards.

have in the different priority queues awaiting upstream transmission. The OLT can
also take into account the service level agreements (SLAs) that have been specified
for the service flows associated with an ONU. The OLT grants bandwidth to each
ONU by sending a GATE message that informs the ONU of the start time and
duration of its transmission on the upstream channel.

4 Comparing the Standards

Both EPON and GPON inherit main standard processes and components from ITU-T
G.983 BPON standard. Examples could be found in their general concepts (PON
operation, ODN framework, wavelength plan, and application). Nonetheless,
significant differences can be found in each modern standard. EPON maintains its
Ethernet-based nature incorporating Ethernet protocols. On the other hand, GPON is
fundamentally a transport protocol that leverages the techniques of Synchronous
Optical NETworking (SONET), Synchronous Digital Hierarchy (SDH) and Generic
Framing Protocol (GFP) to transport Ethernet signals. In this Section we try to
distinguish the major common features and main differences between the two
competent standards. Fig. 5. projects a multilayer comparison approach between the
two standards.

4.1 Layering and Multiplexing

The 10G-EPON is based upon IEEE 802.3 Ethernet that was modified to support P2MP connectivity. It is able to effectively support all Ethernet features. On the other hand, XG-PON constitutes a more flexible paradigm including generic mechanism for any kind of traffic provided. Its adaptation mechanism is quite powerful and is capable of aggregating multi-source traffic stemming from multiple network structures. In 10G-EPON, Ethernet frames keep their original format and properties, offering a flexible and simple layering model. In essence, Ethernet-based PONs carry out IP-based traffic in an end-to-end fashion. XG-PON encompasses Services are all mapped over Ethernet (directly or via IP). This task takes place in XG-PON systems by engaging two layers of encapsulation. Firstly, TDM and Ethernet frames form XGEM frames and then ATM and XGEM frames are both encapsulated into XGTC frames.

In Ethernet-based PONs the Logical Link ID (LLID) addresses an ONU MAC address with OLT ports. This strategy, also known as P2P emulation concept, bridges user and backbone interfaces. On the other hand, XG-PON utilizes a variable called T-CONT to address each ONU. Both LLID and T-CONTs provide a form of P2P emulation.

4.2 Bandwidth Allocation

Reporting the buffer occupancy in the user side is crucial. In the XG-PON, this procedure takes place by inserting it into the fields of the DBRu headers; a single REPORT message is piggybacked in the real data in the 10G-EPON concept. In both cases, the REPORT is mandatory declaring the bandwidth user requests in the uplink direction in terms of Bytes. The transmissions guidelines from the OLT towards the ONUs are provided via the GATE message in Ethernet-based PONs. Usually, a control channel is incorporated to carry out GATE messages. This process is quite different in XG-PON systems. Here, a periodic downstream frame is sent to all ONUs including transmission guidelines for all ONUs in the header. For each ONU the OLT includes the transmission time and the allocated bytes.

In general applying a DBA scheduling algorithm is of paramount importance. However, DBA algorithms are optinal in IEEE 802.3av. The main thinking tank remains the OLT; nevertheless, it's up to the scheduler at the OLT whether / how to construct a transmission schedule. XG-PON employs a very similar scheme, but in that case DBA is part of the standard. Furthermore, the QoS provisioning as a part of the DBA scheme is rigorous and specific.

4.3 Bandwidth and Efficiency

One of the most important assets in the 10G-EPON is the symmetric 10 Gbit/s data rates. Both directions support both , as well as asymmetric 10 Gbit/s downstream and 1 Gbit/s upstream, whereas XG-PON provides 10 Gbit/s downstream and 2 Gbit/s upstream. Efficiency though has to be considered in both directions of a PON. By the term efficiency we usually mean throughput efficiency (also called utilization). Throughput is a measure of how much user data (application-level data) the network

can carry through in a unit of time. Throughput efficiency is a ratio of maximum throughput to the network bit rate. The need for speed in the downstream direction is significantly important while the upstream efficiency guarantees QoS. Each PON protocol introduced its own overhead in either direction. Overall, PON efficiency is a function of protocol encapsulation and scheduling efficiencies.

Apart from throughput, there are two other major parameters for evaluating a PON performance, latency and fairness. The three criteria are interrelated. Reduced utilization results in increased latency because the available bandwidth to empty an ONT queue is decreased; hence more time is required to empty the queue. Similar impact is noticeable for fairness. Low fairness performance indicates that some ONTs will be served slower than others. Consequently, the latency of the slowly-served ONTs increases. Obviously all of the above depend on the DBA algorithm that is being considered (this is considered as future work of the current paper).

4.4 QoS

The QoS provisioning is quite important. Nowadays, the data delivery has been advanced to high-level aggregating multimedia delivery including voice, video, and data capabilities. However, several impairments exist. The QoS concept is realized in the interconnection between the OLT and ONUs. More specifically, the OLT is solely responsible to provide QoS-aware transmission scheduling in both directions without violating the existing SLAs. The PON protocol and architecture provide intercommunication as well as a flow control mechanism that easily facilitates implementation of QoS.

The XG-PON ONU plays a key role in ensuring QoS for all traffic because it is the ingress/egress point for all network traffic. The ONU and therefore all users that are 'hooked' behind it, can experience all types of modern applications. In Ethernet-based PONs this is accomplished by applying strong IP-based QoS support. The ONU can also perform service classification based on the physical port and map it to 802.1p p-bits. For example, traffic flows from voice ports can be classified as highest priority. As part of this service differentiation, the ONU associates different traffic flows with a specific XGEM Port ID. These are virtual port identifiers that have significance for a given XG-PON. In 10G-EPON the QoS provisioning dramatically depends on the allocation mechanism, i.e., the MPCP. There is a large variety of DBA algorithms proposed to use the multipoint control protocol defined in order to arbitrate the transmission of different users. Several DBA algorithms support QoS support in native.

5 Conclusions and Open Issues

The scope of the current paper was to study and compare the two major standards ITU-T XG-PON and IEEE 10GE-EPON for next generation passive optical networks and technologies . The 10G-EPON design aims at exploiting the widespread and mature Ethernet technology for reducing component development effort, design cycles and overall cost. On the other hand, XG-PONs aims at higher line rates accepting higher receiver circuit costs while targeting a set of mechanisms for flexible traffic multiplexing, detailed traffic management specifications and QoS guarantees

with better control of network resource allocation as well as operation and maintenance.

The basic differences of these two technologies were already known from the wide deployment of their predecessors, EPON and GPON. The fundamental difference between 10G-EPON and XG-PON is that XG-PON is a transport technology for Ethernet as well as TDM and ATM. Moreover, GPON utilizes an out-of-band bandwidth allocation map with the concept of traffic containers as the upstream-granted entity. The services are encapsulated into frames in their native format by a process called XGEM while EPON uses no encapsulation (thus, Ethernet traffic is transported natively and all Ethernet features are fully supported). The key aspect of XG-PON's low-latency capability is that all upstream TDMA bursts from all ONUs can occur within 125 µsec. Each downstream frame includes an efficient bandwidth allocation on the BWmap field of the header of XGTC frame, which specifies a bandwidth allocation to a particular Alloc-ID broadcasted to all ONUs and can support a fine granularity of bandwidth allocation. This out-of-band mechanism enables the GPON DBA to support very small grant cycles without compromising bandwidth utilization. XGEM also supports fragmented payloads, which are not allowed in 10G-EPON standard. Thus, a low Class of Service (CoS) T-CONT can stop its upstream burst in the middle of a payload, allow a higher CoS T-CONT its access, and then resume its transmission when told to by the DBA mechanism. Thus, in a highly utilized PON, large bursts of low priority, best-effort data will have minimal affect on high priority, delay-sensitive traffic like voice and TDM. On the upstream direction the two standards work pretty much with the same philosophy. The 10G-EPON uses the MPCP REPORT message and XG-PON the BWmap field in every upstream burst. In that way every ONU lets the OLT know the next desired bandwidth allocation. Of course, the overall merit of the two competitors cannot be judged on technology terms alone. Performance, efficiency and cost have a crucial bearing on deciding which one is the next best thing.

In terms of deployment, although EPON and GPON are widely deployed worldwide, (GPON mostly in North America while EPON in Asia and Europe), there are very few deployments of next generation 10G-EPON and XG-PON. The major reason for this is the relatively high deployment cost for 10G PON which is still in an early state.

Knowing the many benefits of high capacity networks, such as higher bandwidth, enhanced QoS, more bandwidth for each subscriber, lower cost, etc, IEEE points out that will investigate requirements for the next generation of EPON through the activities of the IEEE Standards Association (IEEE-SA). The initiative will attempt to measure the need to support data rates beyond the current top speed of 10 Gbits [17]. Despite deployments of the technology haven't yet taken off, IEEE has decided to test the waters regarding the next set of EPON specifications, as "equipment vendors and network operators, particularly in Asia and North and South America, are interested in exploring the technologies available for the next generation of EPON," according to an IEEE press release [17].

In conclusion it can be seen that both the IEEE and the ITU-T recognized the need to evolve gigabit PONs to 10Gbit/s capable solutions for transporting Ethernet and IP traffic. Due to the need for more bandwidth, they developed 10G-EPON and XG-OPN, respectively. Although, these two standards are very different in execution,

concerning Ethernet transport, management, bandwidth allocation etc, they are equally capable of providing valuable QoS capabilities required for triple play services, each one in a different way.

References

1. Lam, C.F.: Passive Optical Networks: Principles and Practice. Elsevier/Academic Press (2007)
2. Leino, D.: 10 Gigabit-capable Passive Optical Network Transmission Convergence layer design. MSc Thesis, Aalto University (2010)
3. Kani, J., Suzuki, K.-I.: Standardization Trends of Next-generation 10 Gigabit Class Passive Optical Network Systems. NTT Technical Review (2009)
4. Wong, E.: Next-Generation Broadband Access Networks and Technologies. Journal of Lightwave Technology. 30, 597-608 (2012)
5. G.987.1: 10-Gigabit-capable passive optical networks (XG-PON): General requirements. ITU Recommendation (2010)
6. G.987.2: 10-Gigabit-capable passive optical networks (XG-PON): Physical media dependent (PMD) layer specification. ITU Recommendation (2010)
7. G.987.3: 10-Gigabit-capable passive optical networks (XG-PON): Transmission convergence (TC) layer specification. ITU Recommendation (2010)
8. Effenberger, F., Mukai, H., Kani, J.-i., Rasztovits-Wiech, M.: Next-Generation PON—Part III: System Specifications for XG-PON. IEEE Communications Magazine. 47, 58-64 (2009)
9. Cale, I., Salihovi, A., Ivekovic, M.: Gigabit Passive Optical Network – GPON. In: 29th Int. Conf. on Information Technology Interfaces, pp. 679-684 (2007)
10. Ozimkiewicz, J., Ruepp, S., Dittmann, L., Wessing, H., Sylvia, S.: Dynamic bandwidth allocation in GPON networks. In: 4th WSEAS international conference on Circuits, systems, signal and telecommunications, pp. 182-187 (2010)
11. 802.3av: Part 3, Amendment 1: Physical Layer Specifications and Managements parameters for 10 Gb/s Passive Optical Networks. IEEE Standard (2007)
12. Hajduczenia, M., Inacio, P.R.M., da Silva, H.J.A., Freire, M.M., Monteiro, P.P.: 10G EPON Standardization in IEEE 802.3av Project. In: Optical Fiber communication/National Fiber Optic Engineers Conference, pp.1-9 (2008)
13. Luo, Y., Ansari, N.: Bandwidth allocation for multiservice access on EPONs. IEEE Communications Magazine. 43, 16-21 (2005)
14. McGarry, M.P., Maier, M., Reisslein, M.: Ethernet PONs: a survey of dynamic bandwidth allocation (DBA) algorithms. IEEE Communications Magazine. 42, 8-15 (2004)
15. Gaglianello, B.: MPCP Multipoint Control Protocol for EPONs. 802.3ah EPON Security Track (2002)
16. UTStarcom: GEPON Release 2.0 System Overview. Part Number TBD Rev A1 (2007)
17. Hardy, S.: IEEE looks into EPON beyond 10 Gbps. Lightwave (2014)

The Hunter: The Quest for Sensorial Information

Kyriakos K. Skafas[1], Thomas Lagkas[2], and George Eleftherakis[2]

[1] South-East European Research Centre, Thessaloniki, Greece
kyskafas@seerc.org
[2] The University of Sheffield International Faculty, CITY College, Thessaloniki, Greece
{tlagkas, eleftherakis}@city.academic.gr

Abstract. Research is underway in order to identify the challenges posed in the collection, aggregation and transmission of sensorial information from the producers to the consumers over extended, volatile and unreliable networks. Wireless Sensor Networks are comprised of nodes that monitor through sensors interesting targets.
A problem was identified based on a plausible use case, which will hopefully prove to become a base of further research and subsequent development of a satisfying solution. Specifically, the use case involves a small number of sensor nodes compared to a larger number of potential targets. This case taken to an extreme, involves a single mobile node and a single mobile target. In this particular case the ultimate goal of the mobile node is to constantly "hunt" down the mobile target in order to ensure uninterrupted monitoring.

Keywords: wireless sensor network, WSN, mobile ad hoc network, MANET, robot, smart agent, intelligent agent, tracking, received signal strength, RSS, received signal strength indicator, RSSI, triangulation, location

1 Introduction

WSNs are networks of spatially distributed autonomous nodes that monitor various physical conditions via sensors and exchange data wirelessly. Additionally, WSNs are dynamic in nature, due to planned and unplanned changes in their composition or structure. Furthermore, WSNs have a multitude of applications and their applications are expected only to increase in the foreseeable future. Indeed, WSNs are a very active topic of research and development.

There are cases where the relative scarcity of available nodes to the abundance of mobile targets imposes certain mobility and intelligence demands on the nodes. Such cases include the case of a patient in a hospital or clinic, or an athlete on the track or court, or a soldier in the theatre of operations, etc.

In fact, there has always been a need to monitor the status of a patient, especially after a surgical operation and during the subsequent recuperation period. Moreso, concerns are increasing as the world population grows older. Additionally, a shift from re-active management of illness to the pro-active management

of health has been observed even among those members of the population who pose lesser health risks[10, 14].

Ideally, health monitoring would be constant, synchronous and unobtrusive. Such, monitoring would facilitate personalised healthcare, prompt intervention and increased comfort. For example, a patient could wander in the hospital premises and would not be grounded on their bed or constrained in their room, whilst being monitored.

Thus, in order to cover the above cases a system consisting of a Base, a Hunter, and a Prey is described. The Base is the base station, where all the information that originates from a target and is transferred from a node is collected. The Hunter is a robot, that is a mobile node, which includes both sensors and actuators. The Prey is a source of sensorial information, that is a mobile target of interest, which is energy constrained. The Hunter is capable of long range communication with the Base, but follows the Prey, because the Prey is only capable of short range communication.

The most important are the following three:

Detect possible targets in the vicinity.
Select the most interesting target.
Track the selected target, following its movement and maintaining contact.

The rest of this paper is organised as follows: Section 2 provides the problem description, section 3 outlines past work, section 4 specifies out approach to future work and section 5 summarises this paper.

2 Problem Description

The main aspect of this research is the capability of a single node with a single sensor to follow a target irrespective of the type of sensor (e.g. acoustical, optical, electromagnetic, etc.) in order to ensure uninterrupted communication between the target and the base station. As expected, this implies triangulation. Moreover, the need of electromagnetic means of wireless communication between node and target implies, that at least one facility exists on the node that may function as a sensor: a electromagnetic receiver, that may function as a signal strength sensor. This further implies that any other sensor on the node would be superfluous and should be avoided.

Preferably, the node has to surmise where the target has moved, and if it is still moving, then what is its speed and direction, whenever the node detects target movement. In other words, the node should probably be able to move rapidly and randomly in the small area around its initial position and take multiple measurements in rapid succession of the strength of the signal transmitted from the target and received from the node (Received Signal Strenght - RSS) within that area, in order to perform triangulation, whenever it detects an alteration in the strength of signal from the target, as shown in Fig. 1.

Naturally, a plethora of questions arise concerning triangulation. For example, whether triangulation is even possible with a single tracker. It appears that
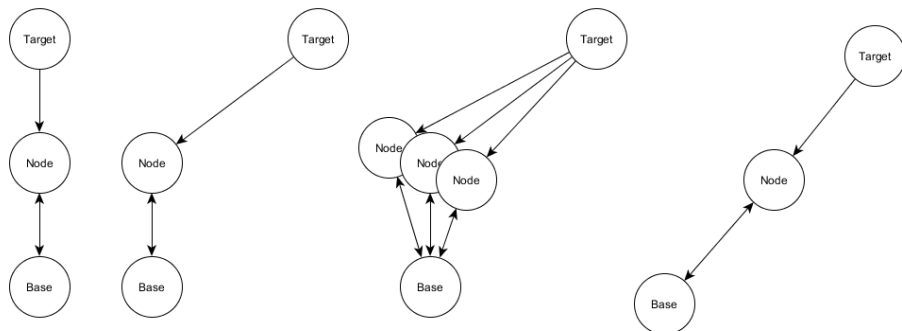
Fig. 1. a. The Initial State, b. The Target Moves: The Node detects the change in the RSS, c. The Node 'dances' around: The Node measures RSS and determines the position of the Target and d. The Final State: The Node repositions itself.

it may be possible, but only if some conditions are met. Namely, that the node ought to be more mobile than the target in order that the node is capable of keeping up with the target's movement and taking multiple spatially distant and temporally close measurements. Also, that the relevant criterion for triangulation is always available, that the necessary triangulation calculations are fast and conclusive, etc. An additional question is whether intermittent target movement affects detection. Certainly, movement affects detection, and thus the node should be able to make take multiple signal strength readings and calculate the position, speed, and bearing of the target near synchronously.

Furthermore, what may be the nature of the algorithm used for the above mentioned operations. There are many choices for each operation, including a generic or simple algorithm, or even a learning algorithm (e.g. a genetic algorithm, an Artificial Neural Network implementation), or maybe well-defined or modeled algorithm (e.g. a closed-loop Control System implementation). The last appears appropriate, since the problem itself is well-defined, and the easiest to implement. Besides, every solution should be adequately flexible and adaptive.

Additionally, a few basic assumptions have to be made in order to simplify the problem and focus on the most interesting parts of a solution. Specifically, that the target continuously broadcasts its sensorial information, which decouples information monitoring, which happens on the target, from information collection, which takes place on the node. Another assumption may be that there is a single node and single target, which further simplifies the problem by merging the aforementioned Detect and Select functionalities of the node. Both target and node move within a flat, two dimensional, unobstructed space, which removes the problem of obstacle avoidance and simplifies triangulation by limiting it to a two dimensional problem. The base is immobile, while both the node and target are mobile, which reduces the degrees of freedom of the system to only four. The node-to-target range is small and the base-to-node range is big, which pushes tracking exclusively to the node. The target may move randomly, but it also

moves slowly, and thus the node is always able to catch up with the target. In addition to the above, there is no signal quality degradation, due to noise or interference, both node and target exhibit no power or energy degradation, and the signal strength to distance correlation is given. Consequently, signal strength is both necessary and sufficient for tracking.

3   Related Work

Location detection may be performed through three main techniques, that is through triangulation, proximity or scene analysis, which may may be used either independently or jointly. Triangulation itself involves lateration and angulation, which use distance and length measurements, and bearing and angle measurements respectively[5].

In detail, lateration requires three distance measurements, each between the point of interest and three other non co-linear points. Measurements may be performed, directly by physically traversing the distance between two points or indirectly by correlating attenuation and distance of a transmission. On the other hand, angulation is similar to lateration with the exception that it also uses angles instead of only lengths. Angulation requires two angle measurements for a two dimensional space and three angles for a three dimensional space. At any rate at least one length measurement is also needed.

Irrespective of the method used, location may be physical or symbolic, relative or absolute, whereas the accuracy and the scale largely depend on the the means to perform the measurements and their precision.

Some systems use the received signal strength (RSS). One such system is RADAR, which was developed by Microsoft[2]. It uses triangulation techniques and IEEE 802.11 wireless networking technologies. It measures at the base stations the signal strength and the signal-to-noise ratio of the transmissions of the mobile nodes and then it uses these measurements to calculate the location of these nodes. RADAR possesses two obvious advantages. Namely, that only a few stations are required and that these stations and nodes are already part of the installed general purpose wireless network.

Other methods, instead of relying on RSS, rely on Time of Arrival (ToA), Time Difference of Arrival (TDoA), Phase of Arrival (PoA), Direction of Arrival (DoA), or Angle of Arrival (AoA) of the received signal[13], while some others combine several[3, 7]. Unfortunately, these methods pre-suppose the availability of lots of processing power or a very precise time source (e.g. ToA, TDoa, PoA), or special receiver with unidirectional antenna (e.g. DoA, AoA), with the former three being less of a issue, due to the recent improvements in microntroller performance, but the latter two posing a valid concern since most widely and readily available wireless products, like those based on IEEE 802.11 (WiFI) and IEEE 802.15.4 (XBee) technologies usually use omnidirectional antennas. In fact, it appears that is the main reason RSS techniques are preferred over others[21].

Moreover, triangulation has also been used extensively in the design of WSNs in order to ensure sufficient coverage. That is, it was used before the implemen-

tation of WSNs in order to achieve optimal node deployment so as not to create overly dense ('fat') or sparse ('thin') areas of coverage[6]. To further elaborate on this point, the strategies of node placement may be placed in three categories, that is as Force Based, as Grid Based and as Computational Geometry Based. All of the various triangulation techniques fall in the latter category and Delaunay Triangulation, which is one such technique, has demonstrated great promise[8, 20, 4].

Additionally, past efforts have demonstrated through the applicability of some techniques in target tracking in addition to node deployment[15, 1]. Still others have qualitatively and quantitatively assessed the relative performance of various such techniques in WSNs either through detailed simulation or extensive experimentation[21, 11].

Also, apart from the numerous methods which exploit RSS actively to estimate the location of targets within the area of coverage of a WSN[12, 9, 16] there are several methods that do so passively[17, 19]. Passive methods present a special interest and are wrought with challenges, because the targets are not active (i.e. not transmitting) or otherwise uncooperative with the nodes and there are no dedicated sensors (e.g. optical sensors) for target detection on the nodes. Thus, with these methods the nodes are trained to interpret the changes they detect in their surrounding environment (i.e. electromagnetic field as perceived through the RSS of their own signals) and then to attribute them with limited certainty (a level of probability) to targets either idle or moving within said environment[18].

Even so, tracking of a mobile target, which is active but uncooperative, by a mobile node, which is single and not part of a larger network of nodes, appears to be relatively uncharted territory.

4   Approach

In order to proceed, the first step to take is to identify an appropriate use case in a case study, and then to describe in detail the requirements implied therein. To that end, research is underway to further elaborate on case a patient wandering within the premises of a hospital or a clinic.

The next step is to make some assumptions in order to focus on the most interesting topics of research. Indeed, some of the necessary assumptions are presented in this paper, but still more could be made e.g. on the effects of path loss (i.e. the reliability and accuracy of the RSS measurements) to triangulation.

After that, the plan is to design the appropriate algorithms and elaborate further with flow charts, block diagrams and pseudo-code when necessary and applicable.

A later step involves producing various simulations using available tools, like NetLogo, MatLab/SimuLink, or Processing. Processing is a language based on Java with which offers effortless development and visual context. Currently, an initial simulation in the Processing programming language was successfully

attempted, that showed great promise, because it pronounced several issues in the original algorithms.

Finally, an actual prototype implementation is to be attempted (preferably using readily available equipment, e.g. a laptop for a base station, a Raspberry Pi for a node, an Arduino for a target, and with WiFi and XBee connectivity). Once an implementation materialises, it may be tested in an experiment occurring in an actual setting and with the various observations considered in and any other feedback incorporated into a future implementation.

5 Conclusion

Currently, focus lies on finding the best way to approach a solution to the one mobile node and one mobile target problem. In other words what should be the actual properties of the Hunter and what the "hunt" truly entails. Obviously, the main limitations of the proposed work derive from the assumptions made in order to facilitate it. Afterwards, future work includes the design, implementation, and testing and verification of the solution.

Once this course of research is concluded and if the development indeed proves successful, then the implication would be that indeed a single node may track a single target successfully, which may lead to further work on node cooperation, emergent node behaviour and node swarms.

References

[1] Alippi, C., Vanini, G.: A rssi-based and calibrated centralized localization technique for wireless sensor networks. In: Proc. IEEE Int. Conference on Pervasive Computing and Communications Workshops (PERCOMW). pp. 301–306 (2006)

[2] Bahl, P., Padmanabhan, V.N.: Radar: An in-building rf-based user location and tracking system. In: INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE. vol. 2, pp. 775–784. Ieee (2000)

[3] Catovic, A., Sahinoglu, Z.: The cramer-rao bounds of hybrid toa/rss and tdoa/rss location estimation schemes. IEEE Communications Letters 8(10), 626–628 (2004)

[4] Chizari, H., Hosseini, M., Poston, T., Razak, S.A., Abdullah, A.H.: Delaunay triangulation as a new coverage measurement method in wireless sensor network. Sensors 11(3), 3163–3176 (2011)

[5] Hightower, J., Borriello, G.: A survey and taxonomy of location systems for ubiquitous computing. IEEE computer 34(8), 57–66 (2001)

[6] Jing, F., Qiong, W., JunFeng, H.: Optimal deployment of wireless mesh sensor networks based on delaunay triangulations. In: Information Networking and Automation (ICINA), 2010 International Conference on. vol. 1, pp. V1–370. IEEE (2010)

[7] Koyuncu, H., Yang, S.H.: A survey of indoor positioning and object locating systems. IJCSNS International Journal of Computer Science and Network Security 10(5), 121–128 (2010)

[8] Kumari, P., Singh, Y.: Delaunay triangulation coverage strategy for wireless sensor networks. In: Computer Communication and Informatics (ICCCI), 2012 International Conference on. pp. 1–5. IEEE (2012)

[9] Kuo, S.P., Tseng, Y.C., Wu, F.J., Lin, C.Y.: A probabilistic signal-strength-based evaluation methodology for sensor network deployment. International Journal of Ad Hoc and Ubiquitous Computing 1(1), 3–12 (2005)

[10] Milenković, A., Otto, C., Jovanov, E.: Wireless sensor networks for personal health monitoring: Issues and an implementation. Computer communications 29(13), 2521–2533 (2006)

[11] Motter, P., Allgayer, R.S., Muller, I., Pereira, C.E., Pignaton de Freitas, E.: Practical issues in wireless sensor network localization systems using received signal strength indication. In: Sensors Applications Symposium (SAS), 2011 IEEE. pp. 227–232. IEEE (2011)

[12] Oka, A., Lampe, L.: Distributed target tracking using signal strength measurements by a wireless sensor network. Selected Areas in Communications, IEEE Journal on 28(7), 1006–1015 (2010)

[13] Pahlavan, K., Li, X., Makela, J.P.: Indoor geolocation science and technology. Communications Magazine, IEEE 40(2), 112–118 (2002)

[14] Pantelopoulos, A., Bourbakis, N.G.: A survey on wearable sensor-based systems for health monitoring and prognosis. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 40(1), 1–12 (2010)

[15] Saxena, M., Gupta, P., Jain, B.N.: Experimental analysis of rssi-based location estimation in wireless sensor networks. In: Communication Systems Software and Middleware and Workshops, 2008. COMSWARE 2008. 3rd International Conference on. pp. 503–510. IEEE (2008)

[16] Su, F., Ren, W., Jin, H.: Localization algorithm based on difference estimation for wireless sensor networks. In: Communication Software and Networks, 2009. ICCSN'09. International Conference on. pp. 499–503. IEEE (2009)

[17] Viani, F., Lizzi, L., Rocca, P., Benedetti, M., Donelli, M., Massa, A.: Object tracking through rssi measurements in wireless sensor networks. Electronics Letters 44(10), 653–654 (2008)

[18] Viani, F., Oliveri, G., Massa, A.: Real-time tracking of transceiver-free objects for homeland security. In: Radar Conference, 2009. EuRAD 2009. European. pp. 621–624. IEEE (2009)

[19] Viani, F., Rocca, P., Benedetti, M., Oliveri, G., Massa, A.: Electromagnetic passive localization and tracking of moving targets in a wsn-infrastructured environment. Inverse Problems 26(7), 074003 (2010)

[20] Vu, C.T., Li, Y.: Delaunay-triangulation based complete coverage in wireless sensor networks. In: Pervasive Computing and Communications, 2009. PerCom 2009. IEEE International Conference on. pp. 1–5. IEEE (2009)

[21] Zanca, G., Zorzi, F., Zanella, A., Zorzi, M.: Experimental comparison of rssi-based localization algorithms for indoor wireless sensor networks. In: Proceedings of the workshop on Real-world wireless sensor networks. pp. 1–5. ACM (2008)

The Role of Contagion in Emotional Multi-Agent Systems

Marina Ntika[1], Petros Kefalas[2], Ilias Sakellariou[3], Ioanna Stamatopoulou[2], and Marian Gheorghe[4]

[1] South East European Research Centre (SEERC)
Research Centre of the International Faculty of
The University of Sheffield, CITY College
24 Proxenou Koromila Str., 54622 Thessaloniki, GREECE
mantika@seerc.org
[2] The University of Sheffield International Faculty, CITY College
13 Tsimiski Str., 54624 Thessaloniki, GREECE
{kefalas,istamatopoulou}@city.academic.gr
[3] Department of Applied Informatics, University of Macedonia,
156 Egnatia Str., 54006 Thessaloniki, GREECE
iliass@uom.edu.gr
[4] Department of Computer Science, University of Sheffield, Regent Court,
Portobello Street, Sheffield S1 4DP, UK
m.gheorghe@sheffield.ac.uk

Abstract. Multi-agent systems have been used to simulate a plethora of scenarios in emergency evacuations. In order to predict human behaviour more accurately, it is important to investigate the effects of incorporating emotions and, in particular, the role of contagion in emotion development. Emotional contagion is the propagation of emotions that is attributed solely to the interaction among individuals. In cases of emergency evacuations, the propagated emotions such as fear and panic, can greatly affect the emergent behaviour of the system. This paper presents three different emotional contagion models and investigates their effects on an open square area evacuation scenario. Additionally, we investigate thoroughly the effects of related model parameters on the simulation. The results support our initial intuitive assumption that the selection of a particular emotion contagion model affects greatly the outcome of the experiments. Different evacuation patterns and different agent behaviours are observed when varying the model, with respect to the numbers of agents that reach various levels of the propagated emotion at different simulation times.

Keywords: Multi-agent systems, simulation, emotional contagion

1 Introduction - Motivation

Multi-agent systems have been used to simulate a plethora of scenarios in emergency evacuations [23,5,12]. In order to predict human behaviour more accurately, it is important to investigate the effects of incorporating emotions and, in

particular, the role of contagion in emotion development. Emotional contagion (EC) is the propagation of emotions that is attributed solely to the interaction among individuals. In cases of emergency evacuations, the propagated emotions such as fear and panic, can greatly affect the emergent behaviour of the system. This paper presents existing emotional contagion models and attempts to investigate their effects on a given evacuation scenario.

A lot of effort has been placed during the past decade on investigating crowd behaviour [24,19,1]. In real-life emergency situations however, people sometimes tend to make irrational decisions, as they are affected by emotions, such as fear and panic. Several researchers attempt to incorporate the effects of emotions on agent behaviour, aiming to develop more realistic simulations [14,18,9]. Towards that goal three factors may be considered of key importance for producing realistic emergency evacuation models and simulations:

- The *emotions theory* on which artificial emotions will be developed;
- The *personality traits* that affect the extend and level of raised emotions;
- The *emotional appraisal* which affect they way the environment is perceived;
- The *communication* which is distorted under specific emotional states;
- The *emotional contagion* which is the propagation of emotion among agents.

Our aim is to investigate the effects of various Emotion Contagion models on the outcome of simulated emergency evacuation scenarios. More specifically, we intend to demonstrate that by maintaining all other parameters unchanged and simply altering the EC model, the experimental results can vary significantly. Additionally, we intend to investigate the factors that may affect the simulation outcomes and identify possible correlations between these factors and the selection of a contagion model.

2 Related Work on Emotion Contagion

It can be intuitively argued that emotion spread acts as an important factor in emergency evacuations, however, to the best of our knowledge, few computational models that deal explicitly with emotion contagion in emergencies exist. The most representative models found in the literature are briefly discussed. Several researchers have attempted to incorporate EC and emotions to various crowd simulations.

Coenen et al. [6] propose an EC model that incorporates a number of *moderating factors*, that are derived from psychological experiments. The EC moderators are: (a) *individual differences* (e.g. gender, personality), (b) *interpersonal factors* (e.g. similarity, group membership), and (c) *miscellaneous* (e.g. pre-existing mood). The resulting simulation investigates elatedness propagation among students and teachers situated in a recreational environment. The authors report that when varying these EC Moderators, the simulation results were consistent with their initial assumptions based on the psychological experiments.

Aydt et al. in [2], propose an emotion engine that incorporates EC. The authors use the appraisal theory [16], according to which emotion is considered

as a cumulative result of the various ways a person assesses their environment with respect to their personal values, goals and well-being. Therefore, in the emotion engine, a set of appraisal patterns is defined and an agent maps the perceived events in its environment to one pattern from the set. Each appraisal pattern is in turn mapped to a specific emotional response. There are two ways to alter the emotion level: either an event triggers emotion intensification or, lacking triggers, the emotion decays with time. The authors use the emotion engine to simulate a protest scenario, and they incorporate anger contagion among the protesters. They also include an authority figure that has a calming effect on the protesters. However, this calming effect is not directly attributed to contagion rather than in worded pacification attempts by the authority figure. An important parameter in this scenario is the anger propagation rate which greatly affects the simulation outcomes.

Tsai et al. in [21] present ESCAPES (Evacuation Simulation with Children, Authorities, Parents, Emotions, and Social comparison) that attempts to simulate crowd behaviour in emergency evacuation. Special focus is placed in the presence of families, since it is usually the case that when family members lose each other amid emergencies, they are capable of even ignoring evacuation directions, and instead focus on locating each other. The authors used ESCAPES to simulate an emergency evacuation scenario in the International Terminal at Los Angeles International Airport, and included an EC parameter in their model to mimic reality more closely. In this scenario, upon interaction between two individual agents, the form of emotion propagation depends on the agent type. For instance, when an agent interacts with a member of the security personnel they inherit the trained officer's lack of fear, and thus experience a calming effect. On the other hand, when an agent interacts with another agent that is panicking, the panic spreads to the former agent as well.

Hoogendoorn et al. in [10] describe contagion as a form of emotion mirroring, and propose a model they call ASCRIBE. For this model, the notion of contagion strength is introduced, which acts as a measure of the influence that the emotional state of an individual has upon another. The main parameters affecting contagion strength, according to the authors, are the *expressiveness* of the emotion sender and the *openness* of the emotion recipient. Bosse et al. in [4] use ASCRIBE to simulate a real event that occurred in Amsterdam, Netherlands, during which, panic was spread in an open area crowd gathering when a person started screaming. This incident resulted in physical injuries to 64 people. Sharpanskykh and Zia in [17] used a variation of the ASCRIBE model to simulate a train station evacuation scenario, where a number of agents are equipped with AmI technology. AmI technology offers to its carriers additional information about the status of clogging of the available exits. Emotion contagion was included, in the sense that when agents interact with AmI equipped ones, they undergo an increase or decrease in their emotional state. The contagion strength depends on a *trust* factor. Trust increases when an agent has a *positive* experience with the information source (i.e. the AmI equipped agent), and decreases on occurrence of *negative* experiences. An experience in this case refers to the

the "verbal" communication of information (e.g. the degree of congestion around an exit) between two interacting agents. The recipient compares the received information against its own beliefs and the experience is subsequently registered as either positive or negative.

Durupinar in [8] propose a different EC model, where interacting agents can be categorized as being either in a *susceptible* or an *infected* state. According to Durupinar's model, a *susceptible* agent may have an evident emotion, and once this emotion exceeds a *threshold* the agent becomes *infected* with this emotion. The *threshold* depends on the individual's empathy, implying that a more empathetic person is rather more prone to becoming *infected*, and has a lower threshold than a less empathetic one. Durupinar discusses a plethora of case studies for this particular model, investigating the propagation of various emotions such as joy (festival scenario), anger (protest scenario) and fear (escape scenario).

3 Design and Implementation of Contagion Models

Inspired by the three basic EC models described, namely ASCRIBE, Durupinar's model and ESCAPES, we have formally defined three corresponding variations of these models [13]:

- C1 mapped to ASCRIBE,
- C2 mapped to Durupinar's model, and
- C3 mapped to ESCAPES.

Our simulated scenario involved an emergency evacuation of an open square area. In this paper, we attempt to describe the models in a diagrammatic way rather than formal, and demonstrate their effect in emergent behaviour on a different setup [13].

3.1 Model C1

In C1, when two agents interact, a series of parameters affect the emotion update for each. The personality affects this update based on the *openness* of the emotion recipient and the *expressiveness* of the emotion sender. The overall contagion strength is additionally affected by the distance between the agents. In Fig. 1, a schematic representation of our C1 model is shown, focusing on one agent acting as the emotion recipient, and the remaining ones are considered as senders. For demonstration purposes we assume in this figure only two marginal levels of openness and expressiveness. In this particular scenario, agent A_1 is assumed to be of high openness, meaning that it is more susceptible to emotional contagion. Both A_2 and A_4 are assumed to be agents of low expressiveness, which can be interpreted as lower capability of emotion passing. The different distances between the interacting agents results in higher contagion strength in the case of A_2 - A_1 interaction compared to A_4 - A_1. On the other hand, contagion strength is higher in the case of A_3 - A_1 interaction, due to the high expressiveness of A_3. Agent A_5 is assumed to be outside the area of influence of agent A_1, and therefore no emotion passing occurs between these agents.
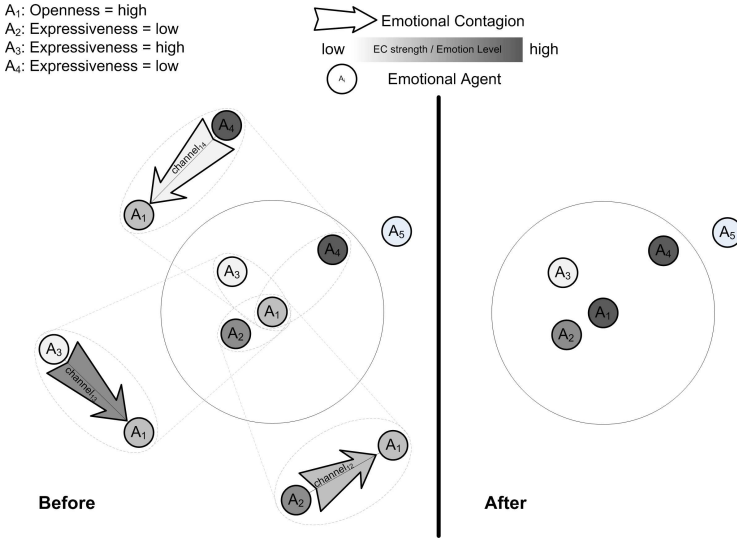
Fig. 1: A schematic representation of model C1

3.2 Model C2

In C2 the agents are categorized as either *infected* by some emotion or *susceptible* to it. When a susceptible individual interacts with an infected one, there is a possibility that the former agent will receive a random emotion dosage from the latter. When the cumulative received dosage exceeds a particular threshold that depends on the agents personality, the susceptible agent becomes infected. The model is schematically represented in Fig. 2.

The figure depicts agents with different values for each of the Big Five [7] basic factors that affect personality (Openness, Consciousness, Extraversion, Agreeableness, Neuroticism). The weights of these parameters on the agent's personality are different for males and females [8]. Depending on these values, the threshold between the susceptible and infected state is defined. In this figure, agents that are susceptible are marked with horizontal lines, whereas infected agents are marked with crossed horizontal/vertical lines. As such, susceptible agent A_1, can only be influenced by infected agents A_2 and A_4. After receiving the corresponding emotion *dosages* d_2 and d_4 from these two agents, it is assumed that the cumulative emotion for A_1 exceeds its predefined threshold, and results in the agent becoming infected as well. In the C2 model, all dosages received by a susceptible agent are cumulated and a total emotion dosage is kept in memory, so once this reaches the threshold value, the transition from susceptible to infected occurs.
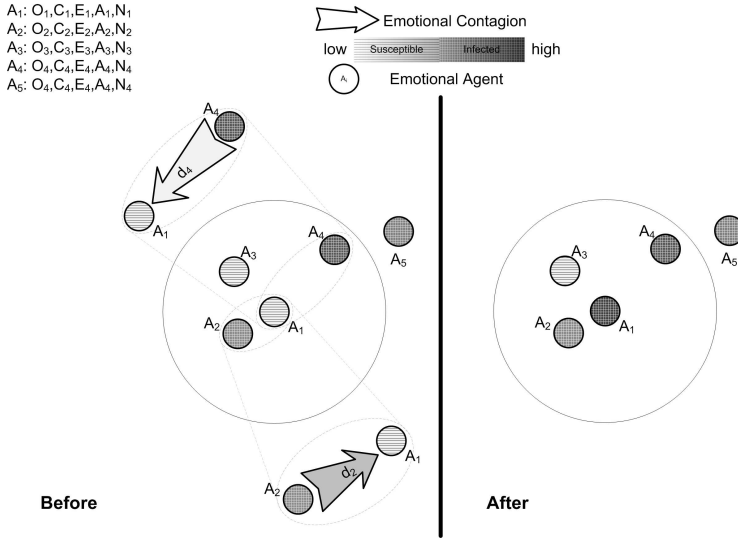
Fig. 2: A schematic representation of model C2

3.3 Model C3

Finally, in our C3 model, when two agents interact, the one with the lower value for the *fear* emotion instantly inherits the higher value of its neighbour. A schematic representation of the model is shown in Fig. 3. Agents A_2, A_3 and A_4 have different values of the same emotion. Since in the figure, the darker shades of grey correspond higher values of emotion, agent A_1 will eventually inherit the emotion level of agent A_4.

All three models were formally described using *emotions X-machines* [11], by including artificial emotion plug-ins within X-machines. The interested reader is referred to [13] for more details.

4 Simulation and Experiments

A simulation for each of our contagion models, C1, C2 and C3 was implemented in NetLogo [22]. It is broadly accepted that the quest for a general approach towards a MAS-based simulation is unattainable. It goes without saying that the inclusion of emotions (with a plethora of different theories for emotions) makes the quest more complex. During the last decade we have invested considerable time to develop executable models for various types of agent architectures and multi-agent systems by building libraries to facilitate those. NetLogo was simple enough to demonstrate the validity of formal models we created and visualise properties that we could not formally verify due to state explosion of model checking. Alternative implementation frameworks were considered at some stages and reviews were published. Certainly, more sophisticated frameworks are now
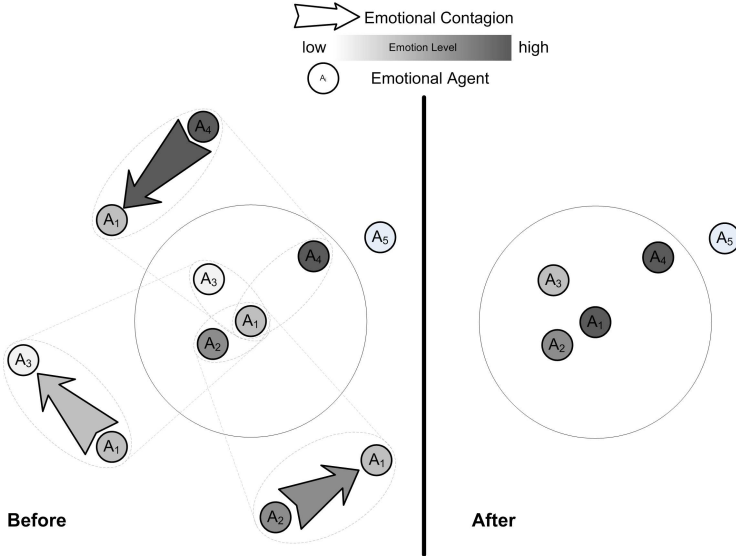
Fig. 3: A schematic representation of model C3

available but we believed that they do not offer many additional features for the purpose we would like to employ them.

The NetLogo model was implemented using a well tested meta-interpreter developed in [15], and that allows direct execution of $^e\mathcal{X}$-machines, that are defined using a simple domain specific language. The meta-interpreter respects the $^e\mathcal{X}$-machines semantics and allows each agent to perform a single transition in each execution step to ensure fairness in the simulation[1].

Some initial experimental results can be found in [13]. These results support our initial hypothesis, that by simply altering the EC model, several differences are observed as the simulation evolves over time. In this paper we significantly extended our experiments to investigate the effect of a number of selected parameters on the simulations' outcome, and identify any emergent important differences between the respective models. Some of the parameters, i.e. number of parents with children, female to male ratio etc., were selected by combining the corresponding factors found separately in ASCRIBE, Durupinar's model and ESCAPES, with the purpose of identifying possible effects of these parameters to all models. More specifically, we compared our three models while varying:

- the total number of agents
- the numbers of authority figures that are present in the area, which are assumed to have a calming effect on evacuees;
- the numbers of family groups among the evacuees;
- the female to male ratio of the population;

---

[1] The implementation can be found at
http://users.uom.gr/~iliass/projects/NetLogo/XMSTATES/

- the vision distance and the danger perception radius for the individuals;
- the effect of an instantaneous source of danger, which is assumed to represent a sudden one-time event, such as an explosion, as opposed to a constant danger source, which could represent continuous events, such as fire.

The propagated emotion in our scenario is *fear*, and the different levels of escalating this emotion for each agent are *Calm, Alarm, Fear, Terror, Panic, Hysteria*.

5 Results and Discussion

The results of altering the number of agents and the number of authority figures in the case of an instantaneous source of alarm have been reported in [13]. Overall, the main conclusion drawn was that the selection of an EC model has a significant impact on the spread of panic, with C1 resulting in the lesser amount of agents reaching *Hysteria* and C3 having the maximum number of agents with this highest emotion level. These findings were consistent throughout all experiments.

5.1 Number of agents

The number of agents appears to influence the emotion spread. As the number of agents increases, in all three models, the corresponding percentage of agents that reach *Hysteria* increases as well. In Fig. 4 this effect is demonstrated for the C1 model. These results are interpreted by considering that the highest the number of agents, the same area becomes more dense in population, resulting in more frequent interactions for each agent. The remaining two models yield similar results, namely the number of *Hysteria* agents rises with an increase to the total number of agents.

The same behaviour emerged in both cases of instant or continuous source of threat. It is worth mentioning that the corresponding graphs for all levels of emotions for all three models were almost identical in both cases.

5.2 Authority figures

The number of authority figures was previously shown to result in a decreased number of agents reaching *Hysteria* in all three models, suggesting that merely the presence of authority figures may help in controlling the emotion spread.

The *calming* effect of authority figures that was observed in experiments for an instant source of alarm is also evident in the cases of a continuous alert. As Fig. 5 shows, as the number of authority personnel increases, the number of agents that reach higher levels of fear decreases dramatically. The graph represents the case of a continuous source of alarm, for the C3 model and for 500 agents, with authority personnel of 0, 20, 50 and 100. Lacking any authority personnel, for C3 model the number of agents that reach *Hysteria* peaks very soon and from there on, it coincides with the total number of agents, meaning that all agents that evacuate are in *Hysteria*. As the number of authority figures rises, the number of agents that remain *Calm* becomes greater.

(a) 500 agents

(b) 1000 agents

(c) 1500 agents
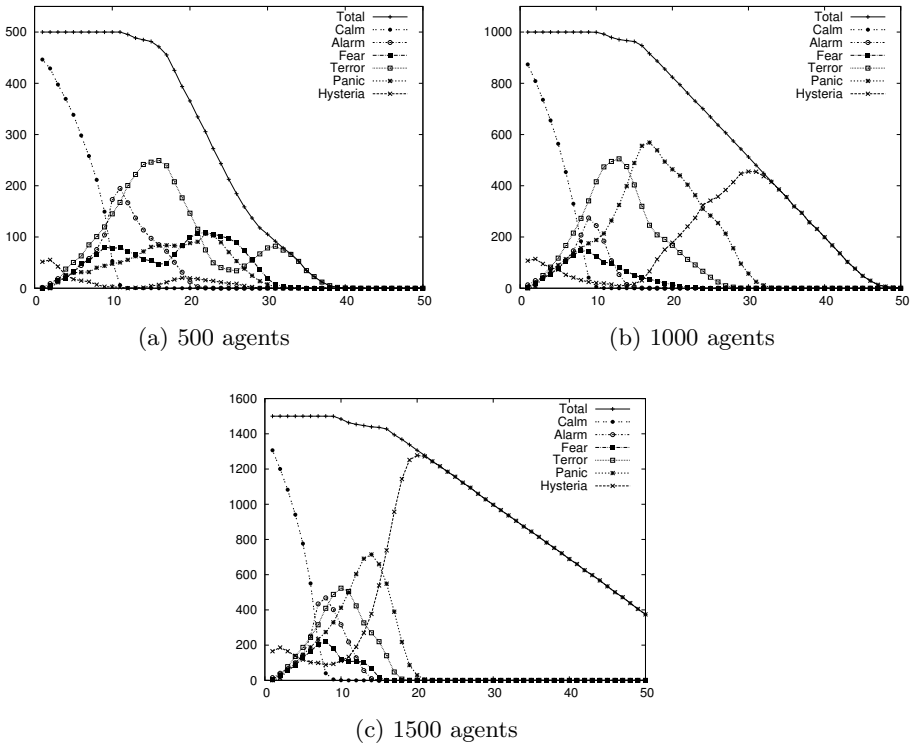
Fig. 4: Model C1 varying the number of agents

5.3 Parents with children

The presence of families with children appears not to affect the simulation outcome for any model. Experiments were ran for 0, 10, 50 and 100 parents with children, out of a total of 500 agents. This was partly anticipated, since our implementation of children-agents did not include any different personality characteristics. What separates them from the other agents is the lower speed with which they can move, and the fact that they are always following their parents. Parents on the other hand, do not evacuate unless their child is safe with them.

5.4 Female to Male ratio

Similarly, we did not observe any significant results from varying the female to male ratio. The only model slightly affected by these variations was C2, as depicted in Fig. 6, where the total number of agents with *Hysteria* is shown. The graph corresponds to experiments ran for the C2 model, with 500 agents and no authority figures. The peak of the curve becomes lower as the percentage of males rises, and a mild shift towards the right, representing a delay in reaching the peak,

(a) 0 authority figures

(b) 20 authority figures

(c) 50 authority figures

(d) 100 authority figures
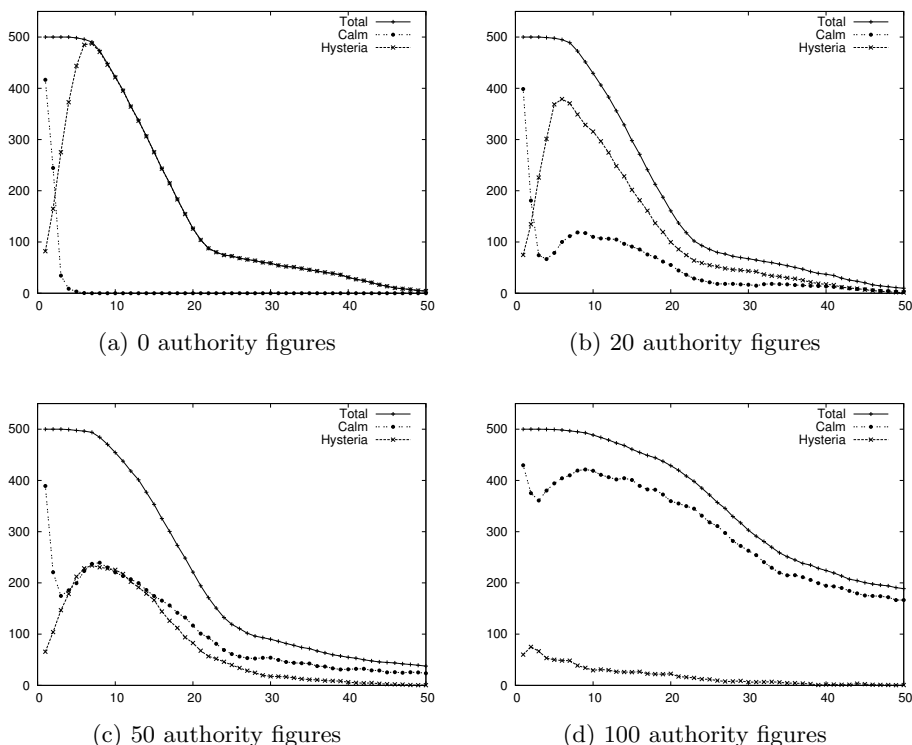
Fig. 5: Model C3 varying the number of authority personnel

can be observed. This behaviour was consistently shown in all experiments with the same model, regardless of the changes in the total number of agents, the number of authority figures and the presence of families with children.

5.5 Perception Distance

A more significant effect was observed when varying the values for the danger perception distance. A large perception distance can be considered as an example of a threat of a greater impact, such as an explosion in the case of instantaneous event or a fire in the case of a continuous event. For all three models, the rise in perception distance correlates with higher number of agents with *Hysteria*. The effects of authority personnel is still evident in the corresponding results. In Fig.7 the C1 model plots, for both an instantaneous and a constant source of danger and for different danger perception distances are shown. Similar patterns were detected for the other two models.

For the cases of small detection distance (which implies that the source of alarm is less obvious to the crowd) and the single and continuous event graphs are
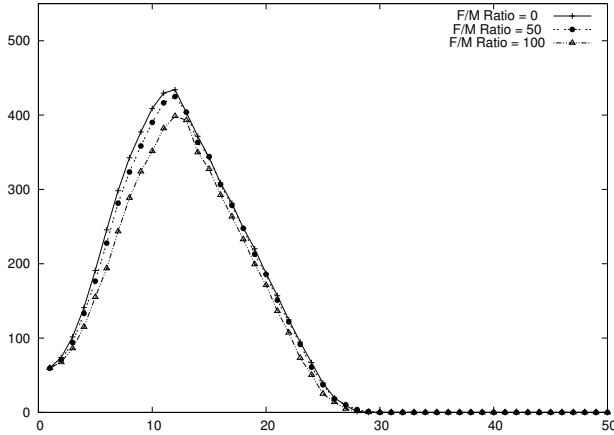
Fig. 6: Number of agents with *Hysteria* in C2 model, varying the female to male ratio, some time after the alarm

almost identical. However, differences emerge as the detection distance increases, leading to an escalation in *Hysteria* for the case of a continuous threat, such as a fire. The observed differences for the two cases of threat duration can be interpreted in a realistic scenario. It is expected that when people are in constant awareness of a threat, their fear levels will rise as the trigger is constantly present. In the case of a single event (a large sound or an explosion), it can be assumed that, lacking the constant stimulation, fear can be more controllable.

5.6   Influence Distance

Another factor that seems to strongly impact the simulation outcome is the influence distance, namely the maximum distance between two interacting agents for contagion to occur. This increase could represent in a real-life scenario the detection of sound in addition to vision. For instance, screaming people can be heard in crowds, even if one has no direct eye contact with them, and this can be assume to induce fear. Bearing this in mind, the results that are shown in Fig.8 describe this scenario quite convincingly.

As the graphs show, for model C2, in the case of a continuous event, and without any authority personnel, for a short influence distance, after about 11 ticks all agents have reached *Hysteria*. When the influence distance is increased to 8, the corresponding peak shifts to almost 3 ticks. However, this is not the case when authority personnel is involved. It appears that, with large numbers of authority personnel, their effect becomes more significant than the emotion spread from other agents. This is also the case when the danger perception distance increases, where the effect of the authority agents becomes less important, but is still significant enough to lead to very controlled emotion propagation.

(a) Singe event - danger perception 10

(b) Continuous event - danger perception 10

(c) Singe event - danger perception 30

(d) Continuous event - danger perception 30

(e) Singe event - danger perception 50

(f) Continuous event - danger perception 50
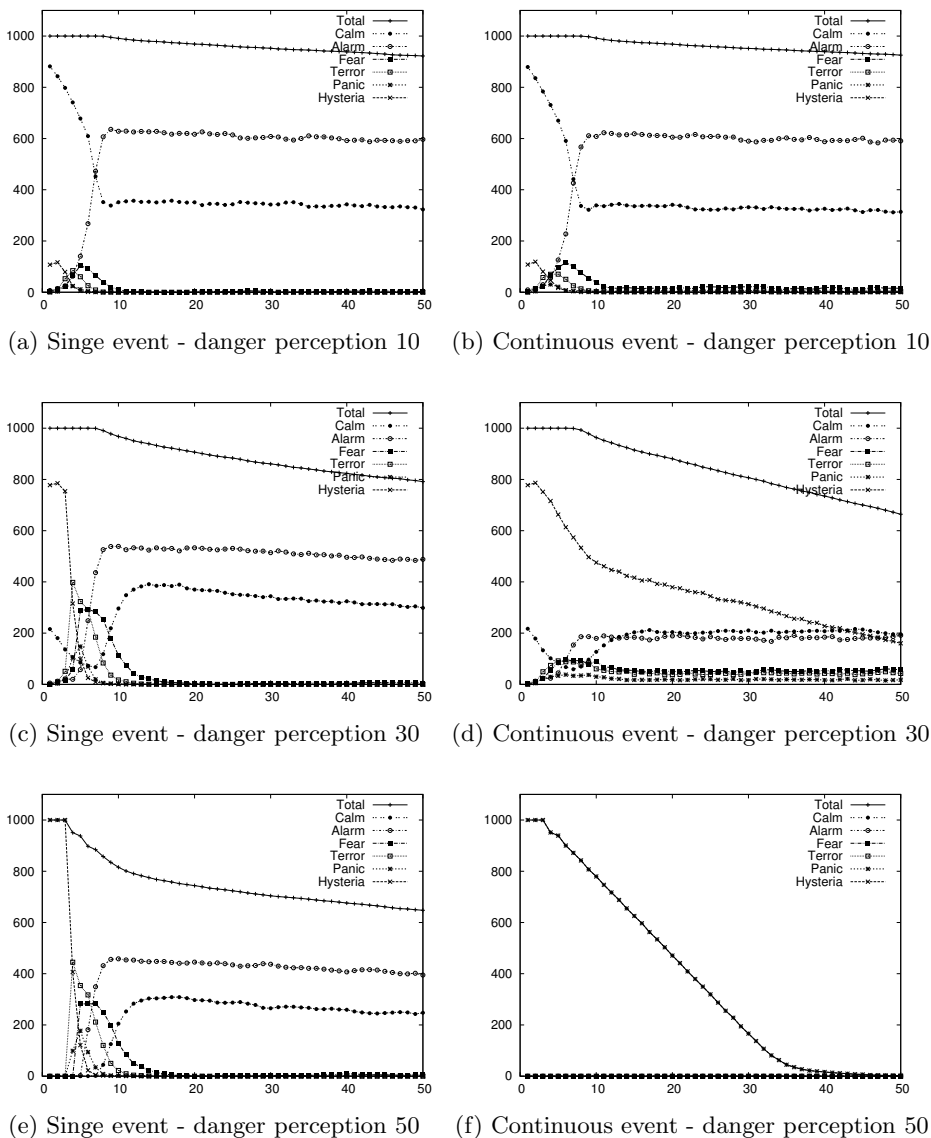
Fig. 7: Model C1, 1000 agents, 50 authority personnel

6 Conclusions and Future Work

After producing the models and running numerous experiments, it can be argued that our initial assumptions are confirmed, in the sense that emotional contagion can be effectively incorporated in simulated evacuation scenarios. The selection
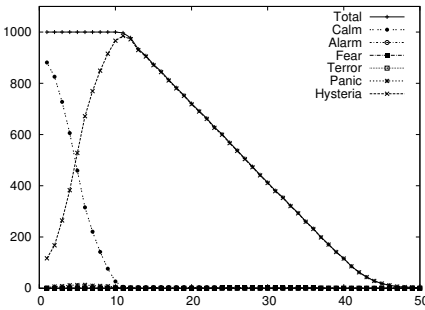
(a) Auth. Personnel 0, Infl. Distance 3
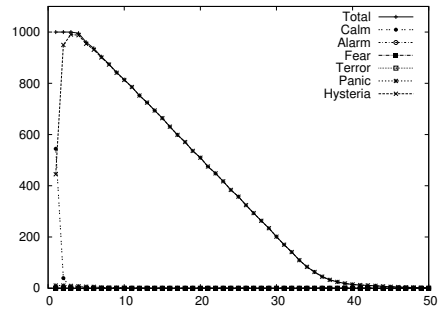
(b) Auth. Personnel 0, Infl. Distance 8
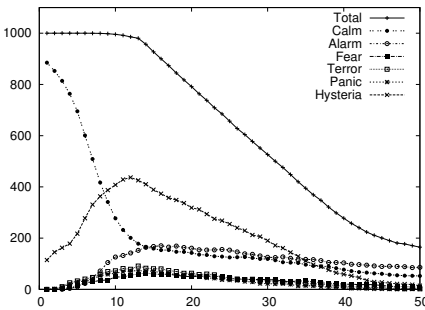
(c) Auth. Personnel 50, Infl. Distance 3
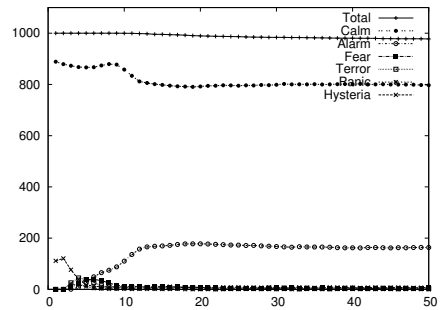
(d) Auth. Personnel 50, Infl. Distance 8

Fig. 8: Model C2, 1000 agents, continuous event, perception distance 10

of the contagion model greatly affects the simulation outcome. Additional parameters that are important in our models have been investigated, and their effects have been thoroughly studied and interpreted.

The main challenge for our future work is calibrating our models and validating them by comparing our results with real life events. So far, our purpose was not to compare our models based on accuracy, since validation requires comparison with real data, which in such cases are scarce, and even when found, they can be rather used for calibration purposes. Instead we aimed to demonstrate that MAS systems can be used to produce simulations that can be characterized as "believable", and we can argue that this has been successfully accomplished.

References

1. Almeida, J.E., Rosseti, R.J., Coelho, A.L.: Crowd simulation modeling applied to emergency and evacuation simulations using multi-agent systems. arXiv preprint arXiv:1303.4692 (2013)
2. Aydt, H., Lees, M., Luo, L., Cai, W., Low, M., Kadirvelen, S.: A Computational Model of Emotions for Agent-Based Crowds in Serious Games. In: Web Intelligence

and Intelligent Agent Technology (WI-IAT), 2011 IEEE/WIC/ACM International Conference on. vol. 2, pp. 72–80 (2011)

3. Bispo, J., Paiva, A.: A model for emotional contagion based on the emotional contagion scale. In: Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on. pp. 1–6 (2009)

4. Bosse, T., Hoogendoorn, M., Klein, M., Treur, J., Wal, C.: Agent-Based Analysis of Patterns in Crowd Behaviour Involving Contagion of Mental States. In: Mehrotra, K., Mohan, C., Oh, J., Varshney, P., Ali, M. (eds.) Modern Approaches in Applied Intelligence, Lecture Notes in Computer Science, vol. 6704, pp. 566–577. Springer Berlin Heidelberg (2011)

5. Braun, A., Musse, S., de Oliveira, L., Bodmann, B.: Modeling individual behaviors in crowd simulation. In: Computer Animation and Social Agents, 2003. 16th International Conference on. pp. 143–148 (May 2003)

6. Coenen, R., Broekens, J.: Modeling Emotional Contagion Based on Experimental Evidence for Moderating Factors. In: Proceedings of the Workshop on Emotional and Empathic Agents in the 11th International Conference on Autonomous Agents and Multiagent Systems. pp. 26–33 (June 2012)

7. Costa, P.T., J., McCrae, R.: Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual (1992)

8. Durupinar, F.: From Audiences to Mobs: Crowd Simulation with Psychological Factors. Ph.D. thesis, Bilkent University, Department of Computer Engineering (July 2010)

9. Helbing, D., Farkas, I.J., Molnar, P., Vicsek, T.: Simulation of pedestrian crowds in normal and evacuation situations. Pedestrian and evacuation dynamics 21, 21–58 (2002)

10. Hoogendoorn, M., Treur, J., Wal, C., Wissen, A.: Modelling the interplay of emotions, beliefs and intentions within collective decision making based on insights from social neuroscience. In: Wong, K., Mendis, B., Bouzerdoum, A. (eds.) Neural Information Processing. Theory and Algorithms, Lecture Notes in Computer Science, vol. 6443, pp. 196–206. Springer Berlin Heidelberg (2010)

11. Kefalas, P., Stamatopoulou, I., Basakos, D.: Formal modelling of agents acting under artificial emotions. In: Proceedings of the Fifth Balkan Conference in Informatics. pp. 40–45. BCI '12, ACM, New York, NY, USA (2012)

12. Murakami, Y., Minami, K., Kawasoe, T., Ishida, T.: Multi-agent simulation for crisis management. In: Knowledge Media Networking, 2002. Proceedings. IEEE Workshop on. pp. 135–139 (2002)

13. Ntika, M., Sakellariou, I., Kefalas, P., Stamatopoulou, I.: Experiments with emotion contagion in emergency evacuation simulation. In: 4th International Conference on Web Intelligence, Mining and Semantics (WIMS'14), 2-4 Jun, 2014 (2014)

14. Pan, X., Han, C., Dauber, K., Law, K.: A multi-agent based framework for the simulation of human and social behaviors during emergency evacuations. AI & SOCIETY 22(2), 113–132 (2007)

15. Sakellariou, I., Kefalas, P., Stamatopoulou, I.: Evacuation simulation through formal emotional agent based modelling. In: Proceedings of the 6th International Conference on Agents and Artificial Intelligence (ICAART 2014). pp. 193–200. SciTePress

16. Scherer, K.R.: Appraisal Theory, pp. 637–663. John Wiley & Sons, Ltd (2005)

17. Sharpanskykh, A., Zia, K.: Grouping Behaviour in AmI-Enabled Crowd Evacuation. In: Novais, P., Preuveneers, D., Corchado, J. (eds.) Ambient Intelligence - Software and Applications, Advances in Intelligent and Soft Computing, vol. 92, pp. 233–240. Springer Berlin Heidelberg (2011)

18. Shendarkar, A., Vasudevan, K., Lee, S., Son, Y.J.: Crowd simulation for emergency response using bdi agent based on virtual reality. In: Proceedings of the 38th Conference on Winter Simulation. pp. 545–553. WSC '06, Winter Simulation Conference (2006)
19. Shi, J., Ren, A., Chen, C.: Agent-based evacuation model of large public buildings under fire conditions. Automation in Construction 18(3), 338 – 347 (2009)
20. Tsai, J., Bowring, E., Marsella, S., Tambe, M.: Empirical evaluation of computational emotional contagion models. In: Vilhjlmsson, H., Kopp, S., Marsella, S., Thrisson, K. (eds.) Intelligent Virtual Agents, Lecture Notes in Computer Science, vol. 6895, pp. 384–397. Springer Berlin Heidelberg (2011)
21. Tsai, J., Fridman, N., Bowring, E., Brown, M., Epstein, S., Kaminka, G.A., Marsella, S., Ogden, A., Rika, I., Sheel, A., Taylor, M.E., Wang, X., Zilka, A., Tambe, M.: ESCAPES: Evacuation Simulation with Children, Authorities, Parents, Emotions, and Social comparison. In: Sonenberg, L., Stone, P., Tumer, K., Yolum, P. (eds.) AAMAS. pp. 457–464. IFAAMAS (2011)
22. Wilensky, U.: NetLogo. `http://ccl.northwestern.edu/netlogo/` (1999), center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL
23. Zaharia, M.H., Leon, F., Pal, C., Pagu, G., Baykara, N., Mastorakis, N.: Agent-based simulation of crowd evacuation behavior. In: WSEAS International Conference. Proceedings. Mathematics and Computers in Science and Engineering. No. 11, World Scientific and Engineering Academy and Society (2009)
24. Zarboutis, N., Marmaras, N.: Design of formative evacuation plans using agent-based simulation. Safety Science 45(9), 920 – 940 (2007)

Bind server – an adapter from high level to embedded communication in service oriented architectures

Kujtim Hyseni[1] and Neki Frasheri[2]

[1] Department of Information Engineering, Faculty of Information Technology,
Polytechnic University of Tirana, Mother Teresa Square 1, Tirana, Albania
[2] Center for Research and Development in IT, Faculty of Information Technology,
Polytechnic University of Tirana, Mother Teresa Square 4, Tirana, Albania
[1]kujtimhyseni@hotmail.com, [2]nfrasheri@fti.edu.al

Abstract. In this paper is presented the process of generating code which has the objective the integration of 8051 microcontrollers as primitive embedded devices to conventional service oriented architectures (SOA). We base our study in CORBA distributed platform. Microcontroller transfers and applies requests/responses between conventional machines and sensors/actuators. Since it is impossible to implement SOA technologies in primitive devices such as 8051 microcontrollers a solution is presented which introduces bind servers to mediate between high level SOA requests to low level intended for microcontroller. We present the automated generation of code for bind server basing on Interface Definition Language (IDL) as input for JacORB – a Java implementation of CORBA.

Keywords: embedded system, service oriented architecture, code generation, CORBA.

1   Introduction

Microcontroller is an embedded device with on-chip consisting CPU, RAM and ROM. Amounts of RAM and ROM are small while CPU supports instructions (commonly 8 bit) limited in capabilities. Typical on-chip RAM consists of 256 bytes, while ROM of 12 kbytes [1]. Although limited in power, they are inevitable back-end components in configuration with more powerful devices such are personal computers (PCs). In data acquisition and control systems, they typically serve for translation of higher (byte) level requests coming from PC to lower (bit) level intended to given sensor or actuator. This is possible due to its pin bit-addressable nature. To each bit of byte corresponds a logical (voltage) level at microcontroller specified pin or pins. PC hosts a database and application which feeds the database with data acquired from sensors with the help of microcontrollers. Also, in similar way an actuator may be driven from application via microcontroller. Works [2],[3] employ microcontroller for data acquisition and control purpose. But, these works employ it in ad-hoc way i.e. based upon given problem nature. This approach has an advantage since its resources such as pin connections are used optimally depending on the concrete problem. In

these configurations microcontroller is connected to PC and waits requests from it. This way, it is considered isolated being accessible only by single PC through its common interface which in most cases is serial port. An advantage would be making possible to access the microcontroller through network as standard network component. It would be suitable to access it as a component of some of standard Service Oriented Architectures (SOA) such as CORBA [4] or Web Services [5]. In other words it should be considered as a CORBA or Web Services server which reacts on client's requests. However, due to its limited capabilities and very limited resources it is impossible to implement above mentioned technologies mechanisms in microcontroller, so we have to look for other solutions. The implementation of technologies should be done in automatic way, possibly with the help of tools known as code generators. This would avoid the errors which may happen during manual work as a result of inattention of the developer.

2 Related work

A number of works deal with the problem of implementing SOA technologies in embedded devices.

Van Engelen [7] introduces gSOAP development environment, specialized for code generation and runtime optimizations for developing light-weight XML Web Services for embedded devices in C and C++. Main drawback of this work is that for generated code, although optimized, doesn't exists a compiler for target embedded device with support for yet complex C and C++ libraries. Moreover, the generated code doesn't compiles with popular C compiler namely SDCC [8] for 8051 [6] microcontroller even some properties such as Input-Output are excluded. By Input-Output is meant the communication with user which commonly is the console or other standard mechanism. The other drawback is that it uses standard XML-based Web Services protocols that are inconvenient to implement in restricted environments.

Bagnasco et al. [9] apply a common Web Service paradigm to both small devices and Internet network nodes in order to promote systems interoperability and scalability. They present two layered service oriented architecture. In the lower layer reside devices performing basic tasks, such environmental sensing and conditioning. Lower layer is REST based. It interfaces with high layer nodes that offer more complex services on the base of the collected results. Upper layer is SOAP (Simple Object Application Protocol) based, and is exposed to external world such as internet as standard Web Service.

More recent and interesting works are presented by Käbisch et al. in [10], [11] and [12]. These papers are of particular interest since they employ microcontrollers – although more powerful than we do. While in [10] is presented the work in general, showing Web Service code generation and communication based on the EXI (Efficient XML Interchange) [13] format, [11] and [12] show further improvements. According to [10] the automated workflow of Web Service generator for microcontrollers is mainly divided into two phases. In first phase, based on a given service description, schema information is generated that describe all possible messages which can be understood by the service. In second phase, source codes are generated which contain, the EXI Processor based on schema generated in first phase,

RPC (Remote Procedure Call) skeletons, and a dispatcher that handles the request/response messages. Paper [11] takes the advantage of EXI specification [13] on transforming schema information to EXI grammars. The reason of such a transformation is that EXI grammars are much simpler to process, compared to the equivalent XML schema information. EXI grammars correspond to deterministic finite automata (DFA). Given that XML is not a regular language, a single EXI grammar cannot be used to represent an entire XML document. Instead, an EXI coder uses a stack of grammars, one for each element content model (e.g. Envelope Grammar). The underlying Document Grammar stack item in turn describes root elements that may occur in an EXI document. EXI results in a low message sizes which are suitable for transmission but there is an overhead in processing i.e. compression and decompression which results in a considerable code size.

Taking into consideration drawbacks of above works we propose alternative solution. It is clear that SOA architectures cannot be implemented in primitive devices we target. Thus, there is a need to attach powerful device to primitive embedded one such will obtain complex tasks, while the embedded device will obtain simple tasks such as the communication with sensor or actuator. The powerful device would run the corresponding ordinary server (CORBA server or Web Service server) while the embedded device a simplified version of it. The server hosted in powerful device would translate complex requests from coming client to simple ones dedicated to embedded device and vice versa, responds from embedded devices back to client. This paper deals with generation of such server which mediates between powerful device such as PC and embedded device.

3 Architecture and approach presentation

3.1 Microcontroller as component of Service Oriented Architecture (SOA)

Since we deal with very primitive devices such as 8 bit microcontrollers neither of described works in section 2 can be implemented. That is because they employ large code footprints and high memory resources. Another difficulty is that there is no complier available for 8051 architecture for the code generated say with tools [7]. Considering these statements, we saw the need for different approach to the problem. Instead, we propose the sharing of tasks between the microcontroller and conventional machine (PC) attached to it. Since PC is more powerful it would obtain complex tasks while microcontroller more simple tasks. It would be useful to see the microcontroller from the perspective of the client as a standard CORBA or Web Service server. This way, the PC would host say the CORBA server accessible from client, while the same server would bind to the connected (attached) 8051 server and translate complex requests coming from the client to simple dedicated to 8051 server running in microcontroller. So, microcontroller would run simplified version of CORBA server. We went more far, such we designed platform independent to technology such that the microcontroller server is accessible from various SOA technologies at the same time. Our solution is to attach a bind server that is specific for the given technology. The server specific to technology would translate requests

specific for architecture to platform independent 8051 server understandable requests (figure 1).
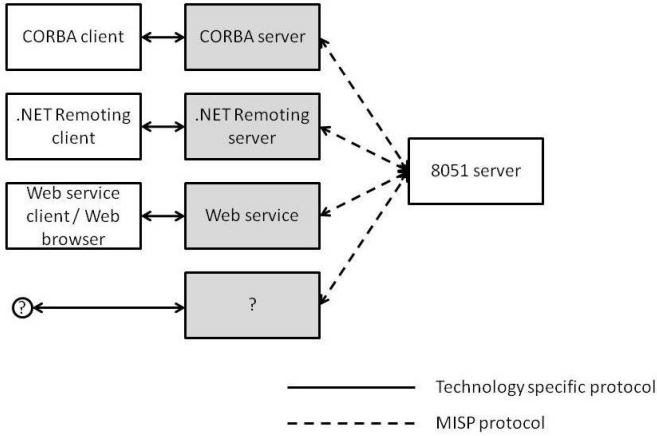
Figure 1. Architecture independent 8051 platform

As a basis for this approach lays the MISP (Minimal Inter-Server Protocol) protocol [20] (figure 2), which serves for communication with 8051 server. It is technology architecture independent, and is programming language oriented. As can be seen from figure 2 it allows to address the server (in CORBA the IDL interface), the operation and operation parameters. The question is put on how to support attributes in CORBA IDL definition. With little effort they can be easily implemented with the MISP. Attributes can be set or get. Both of these actions can be considered as ordinary operations. Get would be an operation that returns value of given attribute type, while set, would be an operation with input parameter of given attribute type. Since by one byte each is dedicated for server and operation addressing, their total count number of declaration is limited to 256. For server addressing (interface), it is obviously enough but it suffices also for operations allowed to be declared. Considering that microcontroller serves to obtain single or few tasks, the total number of operations including attributes can be a few.

1 byte	1 byte	1 byte	n bytes	m bytes	...

Msg length	Server ID	Operation ID	Param 1	Param 2	...
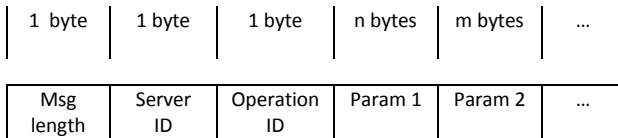
Figure 2. MISP message fragmentation components

3.2 Implementation through automated code generation

Developing servers manually even if there is template used is tedious and error prone. Thus, we have looked for possibilities to obtain these tasks with the help of tools known as code generators. Only the communication of bind server with 8051 server

has to be generated. Bind server is to be completely generated and should not be changed by the developer since it is optimized for given architecture. What has to be changed/appended by the developer is the implementation of operations in 8051 server. Operations are application specific. Commonly they deal with communication with given sensor(s) or actuator(s).

We did the implementation of code generator for CORBA architecture communication. As CORBA implementation is chosen JacORB [17], a Java implementation of CORBA. Thus the bind server is to be generated in Java, while the 8051 server currently is generated in 8051 Assembler [19]. Again, the communication between the bind server and 8051 server is obtained through the MISP protocol which is technology independent. Our code generator deals with generation of that communication. Although there are available specifications for embedded devices such as CORBA/e [15], these specifications result in large code sizes and high memory bandwidth to be implemented in say 8 bit 8051 microcontroller with only 12 kilobytes of flash program memory and 256 bytes of RAM [1]. Another requirement is that Java Virtual Machine (JVM) has to be run in microcontroller where the server has to be hosted and run. At the moment when this paper is written there is no JVM available for 8051 microcontroller. For these reasons we simplified the problem such microcontroller 8051 will support only simple IDL constructs. These include single interface declarations with operations and attributes of basic types excluding *string* and *wstring*. The grammar of acceptable IDL file is described by only 14 rules by contrast to recent IDL specification [14] which contains a total of 138 rules. Grammar of our code generator is shown in listing 1 below.

Listing 1. Coco/R acceptable IDL definable reduced grammar rules

```
MISPIDL = interfacedef.
interfacedef = "interface" identifier "{" interfacebody
"}" ";".
interfacebody = {interfacelement}.
interfacelement = methoddef | attributedef.
methoddef = onewaymethod | standardmethod.
onewaymethod = "oneway" "void" identifier "("
[onewaymethodparammetersbody] ")" ";".
standardmethod = simpletypesplusvoid identifier "("
[methodparammetersbody] ")" ";".
simpletypes = "octet" | "char" | "boolean" | "short" |
"float" | "double" | "long" ["long" | "double"] |
"unsigned" ( "short" | "long" ["long"]).
simpletypesplusvoid = simpletypes | "void".
Onewaymethodparammetersbody = onewaymethodparammeter {
"," onewaymethodparammeter}.
methodparammetersbody = methodparammeter { ","
methodparammeter}.
onewaymethodparammeter = "in" simpletypes identifier.
methodparammeter = ("in" | "out" | "inout") simpletypes
identifier.
```

```
attributedef = ["readonly"] "attribute" simpletypes
identifier { "," identifier } ";".
```

This grammar is transformed such that is acceptable by Coco/R code generator tools. Coco/R [16] is a compiler generator, which takes an attributed grammar of a source language and generates a scanner and a parser for this language. The scanner works as a deterministic finite automaton. The parser uses recursive descent. LL(1) conflicts can be resolved by a multi-symbol lookahead or by semantic checks. We had to supply the main class that calls the parser as well as semantic classes that are used by semantic actions in the parser. A semantic action is a piece of code that is written in the target language of Coco/R (in Java in our case) and is executed by the generated parser at its position in the production.

4   Bind Server generation

Bind server, as the name suggests, serves to mediate between the client and the server implemented in 8051 microcontroller. It is realized as CORBA server for JacORB. Thus to the client it appears and is accessible as common CORBA server. Intrinsically, it detects which operation is invoked and builds stream conform MISP protocol intended for 8051 server. The stream is built such that code generator calculates its size basing on parameter types sizes and adds three bytes for message header. Message header consists message length, server ID (which by default is 0) and operation ID. Other bytes serialized are *in* and *inout* parameters as appear in order from left to right in operation declaration in IDL file. IDL file has to be written according the grammar rules shown in Listing 1.

Bind server is completely generated by the code generator tool we built and accepts the grammar shown in Listing 1. It should not be changed since it conforms the strict rules of message exchange with microcontroller.

4.1   Bind Server details

Bind server class definition extends the POA (Portable Object Adapter) class generated by JacORB code generator tools. Further, it employs two other classes, *MISPSerial* and *StreamOperations* classes to run properly. *MISPSerial* is responsible for forwarding the serialized (streamed) requests to standard integrated 8051 COM port. It is bound to the port and sends requests towards the microcontroller and waits for the responses from microcontroller. Since the stream may come in groups from microcontroller, for instance, due to use of non-standard USB-to-COM converters, it takes care of proper building the incoming stream. For port operations it uses rxtx [18] free open source project Java libraries. *StreamOperations* class is responsible for building the stream, creating stream header, serializing the operations parameters and de-serializing them back to their actual values.

The communication with microcontroller is strictly client-server, meaning that microcontroller is strictly a server responding to requests initiated only by bind server. Another characteristic is that communication is synchronous, meaning that only one

request is processed and responded at given time. When more operations are invoked during the one is processed, they are added to queue by bind server through its inherited JacORB mechanisms and will be executed in order as they arrive. The first arrived is executed first and last arrived is executed the last.

At beginning of bind server class code are generated two constructors, one for default values for port name and baud rate and one to be specified by the user. Next, are generated operations in the order as they appear in the IDL file. If there are attributes, corresponding operations are generated to *get* and *set* respectively the value of attribute. If there is *readonly* attribute, only the operation for *get* is generated. At each operation the operation ID is generated in incremental order from first to the last. Output and input stream sizes are calculated by code generator and they are to be used by bind server to specify the output stream to be sent and to check the size of the received input stream size. Then, the request (out) stream is created and set the values of size, server ID and operation ID each by one byte. Next, *in* and *inout* parameter values are written to the stream with the help of *StreamOperations* class. The parameters are written in the order as they appear in operations declaration in IDL file. Then, the stream is sent to the microcontroller through its built in serial port. The loop for waiting the response from microcontroller is written to wait for complete receiving of the stream by microcontroller. The loop employs the MISP class to take care to receive the complete stream even if it comes in groups. After receiving the response from microcontroller, the size of the stream, the server ID and operation ID is checked to ensure that the expected stream is received. Upon successful check, the values to be returned by operation (if it returns non-void), *out* and *inout* parameter values are read from stream and set to corresponding parameters. Upon receiving the wrong stream the user is informed.

4.2 Bind server generation and running procedure

The first thing to be done is the writing of IDL file which defines the functionalities the microcontroller provides. Then, with the help of JacORB code generation tools are generated the required stubs and skeletons which allow the requests/responses to be called and transferred over the network. Following, with the help of our code generator tools are generated the bind server and microcontroller server file. Bind server file is complete and should not be changed, while microcontroller file should be appended with functionalities the microcontroller provides. Commonly, this is the communication with given senor and/or actuator. Then, the bind server is compiled with JacORB tools as ordinary CORBA server. After successful compilation, the server is run. At this stage, it is ready and waits for requests coming from client which are to be translated to MISP compliant acceptable by 8051 microcontroller and back from MISP compliant to those CORBA compliant acceptable by client.

5 Conclusions

In this paper is presented the automated way of generation of bind server code. Bind server mediates between and adapts high level client requests to simple dedicated to

primitive embedded 8051 microcontroller, and vice versa, low level responses coming from embedded microcontroller towards high level dedicated to client. The implementation is done for CORBA platform, but, the generated server with few code lines modification can be transformed to Web Service server. This is possible since our approach is language oriented – the generated code is related to given programming language for which the SOA technology exists, in our case Java.

6 Future work

Since we aim to support multiple technologies for microcontroller i.e. to allow the microcontroller to be accessed by various clients implemented in other technologies than CORBA say Web Services the connection with microcontroller via its serial port should be available during most of the time. Thus, one of the main contributions to future work is the problem of managing the connection with microcontroller. But, this is not always suitable and even possible. This is because the period of invoking operations should be greater or in worst case equal to the times of establishing the connection with microcontroller plus the delays in network transport. So, if we aim at managing the connection smartly by bind server the above times should be considered. Obviously, an initialization time measurement process is required to choose the connection policy by bind server. As other future work is the implementation of error throwing by microcontroller which has to be managed by bind server too. Next work would be the support for building dynamically the protocol during code generation phase. Building the protocol dynamically is of importance due to the fact that optimized communication with microcontroller server is of particular interest due to its limited memory resources.

References

1. Atmel Corporation: AT89S8253 Datasheet. (2008)
2. Szarek, T. R.: On the use of microcontrollers for data acquisition in an introductory measurements course. Master's Thesis, University of Notre Dame, Indiana (2003)
3. Saraf, S. R., Holmukhe, R. M.: Microcontroller based data acquisition system for electrical motor vibrations using VB software. Indian Journal of Computer Science and Engineering (IJCSE) (2011)
4. OMG's CORBA Web-site, http://www.corba.org/, Last accessed on 21 January 2014
5. W3C Working Group Note: Web Services Architecture., http://www.w3.org/TR/ws-arch/ (2004) , Last accessed on 21 January 2014
6. MacKenzie, I. S.: The 8051 Microcontroller. Prentice Hall, Third Edition (1995)
7. Engelen, R. V.: Code Generation Techniques for Developing Light-Weight XML Web Services for Embedded Devices. Proceedings of the ACM SIGAPP SAC Conference (2004)
8. SDCC - Small Device C Compiler, http://sdcc.sourceforge.net/, Last accessed on 16 January 2014
9. Bagnasco, A., Cipolla, D., Occhipinti, D., Preziosi, A., Scapolla, A. M.: Application of web services to heterogeneous networks of small devices. In: Proceedings of the 8th WSEAS Int. Conference on Automatic Control, Modeling and Simulation, Prague, Czech Republic (2006)

10. Käbisch, S., Peintner D., Heuer, J., Kosch, H.: XML-based Web Service Generation for Microcontroller-based Sensor Actor Networks. In: 8th IEEE International Workshop on Factory Communication Systems (WFCS) (2010)
11. Käbisch, S., Peintner D., Heuer, J., Kosch, H.: Efficient and Flexible XML-based Data-Exchange in Microcontroller-based Sensor Actor Networks. In: IEEE 24th International Conference on Advanced Information Networking and Applications Workshops (2010)
12. Käbisch, S., Peintner D., Heuer, J., Kosch, H.: Optimized XML-based Web Service Generation for Service Communication in Restricted Embedded Environments. In: IEEE 16th Conference on Emerging Technologies & Factory Automation (ETFA) (2011)
13. Schneider, J., Kamiya, T.: Efficient XML Interchange (EXI) Format 1.0. W3C Candidate Recommendation 10 March 2011 (2011), http://www.w3.org/TR/exi/, Last accessed on 16 January 2014
14. Common Object Request Broker Architecture (CORBA) Specification, Version 3.3, Part 1: CORBA Interfaces, http://www.omg.org/spec/CORBA/3.3/Interfaces/PDF, Last accessed on 16 January 2014
15. Common Object Request Broker Architecture (CORBA) for embedded (CORBAe), http://www.omg.org/spec/CORBAe/, Last accessed on 21 January 2014
16. Mössenböck, H.: The Compiler Generator Coco/R - User Manual. Johannes Kepler University Linz, Institute of System Software, Linz (2010), http://www.ssw.uni-linz.ac.at/coco/, Last accessed on 16 January 2014
17. JacORB - The free Java implementation of the OMG's CORBA standard, http://www.jacorb.org/, Last accessed on 16 January 2014
18. RXTX for Java, http://mfizz.com/oss/rxtx-for-java, Last accessed on 12 March 2014
19. 8051 Macro Assembler ASEM-51, http://plit.de/asem-51/, Last accessed on 13 March 2014
20. Hyseni, K., Frasheri, N.: Microcontrollers as components of service oriented architectures. In: 3rd Mediterranean Conference on Embedded Computing (MECO) (2014)

An Investigation into the Use of Abstraction in Model Checking Z Specifications

Maria Siregar[1], John Derrick[2],

Verification and Testing Lab., Dep. of Computer Science, The University of Sheffield, UK
[1]acp12mus@sheffield.ac.uk, [2]j.derrick@dcs.shef.ac.uk

Abstract. Z notation is a language used for writing formal specifications of a system. However, there is less tool support for this language. One such tool that is not generally available is a model checker. Model checking is a method used to verify that a system has certain properties, and has been found to be useful since it can provide full verification of a finite state system without the user having sophisticated knowledge. Originally applied in hardware systems, it is now commonly available for application in software systems. One of the drawbacks of model checking is that it applies to finite state systems, since it works by performing a complete state space exploration. However, the size of the systems that model checkers can now cope with has increased rapidly. In this paper, it is investigated the use of abstraction as a means to make model checking feasible for arbitrary Z specifications.

Keywords: Abstraction, Z specification, Z2SAL, SAL model checker, model checking.

1 Introduction

We are in an era surrounded by computer applications. Although those applications are used in almost every aspect of our life, we always need them to do their jobs accurately, particularly safety-critical systems.

Directed by that need, previously, natural language and graphics were used to draw systems flowcharts and to write specifications. However, natural language is inadequate for writing such specifications due to its imprecision. The alternative, which is the use of a programming language to write a specification, is equally flawed in that it forces one to work at the wrong level of abstract [1].

The method of writing specification should be precise enough as well as implementation free. Moreover, such a method, if it is equipped with a proof theory, can help us to describe properties of specifications easily by conducting 'rigorous arguments' [1]. It raises a need for a certain level of formality and for specifications to be written at a suitably high level of abstraction. Thus, mathematical notation is used which is based on set theory, logic, functions and relations to write those specifications. Notations used to do this are called specification languages or *formal methods*. Indeed, although their use is not widespread in every sphere, formal

methods are recommended by many standards bodies concerned with Safety-Critical systems [2].

Z is one example of formal language. It can be used to write a precise and ambiguity free specification. It also can be analyzed mechanically [3].

However, the increasing interest in the use of Z does not balance with the amount of available tool support for Z. There are many aspects to this situation, such as the abstraction and logic of that language is undecided [3].

As the logic of Z notation cannot be decided that makes even the simple theorem hard to prove, model checking seems to be the answer for verifying Z specification [3]. Model checking is a verification method that is an automatic, model-based, and property-verification approach. It relates closely to concurrent or reactive systems, and stemmed from a post-development methodology [4]. One does not need an expertise in mathematical disciplines to be able to model check such systems [5].

Nevertheless, there is hardly any tool support for model checking Z specifications. Although the *Community Z Tools* (CZT) project is developing continuously a set of open source tools for Z, its progress is slow [6]. In particular there is currently no model checker available for Z either in CZT or elsewhere.

Furthermore, if such a model checker were to exist for a Z specification, it has drawbacks. The two principle ones being that it only applies to finite state systems, and even then these cannot be too large since it can suffer from *state space explosion* problems [5, 7, and 8]. Such an explosion has been shown to be the most challenging problem on model checking [5, 7, and 8].

Smith and Winter [9] reported that a Z specification could consist of complex predicates as well as large number of or even infinite state spaces. As a result, researchers such as Jackson, Smith and Winter [3, 9] have proposed the approach of abstraction to Z specification.

In this paper, it will investigate the use of abstraction in model checking Z specifications to enable a model checker to verify various Z specifications. This paper also investigates to what extent that abstraction process could be automated. This work refers to the abstraction process of Smith and Winter [9].

The structure of this report is as follows. Section 2 describes some research related with this paper. Followed by is Section 3. This section provides the sources that relates to verifying Z specifications. The Next section relates with abstraction on Z specification. Future work and conclusion is given in Section 5. Acknowledgements are given at the end just before references.

2 Related Works

Here only works that are related to abstraction on Z specifications will be described. There are two papers published on that research. It will start with a paper of Jackson [3].

Jackson proposed model checking as a method of analyzing Z specification [3]. Not only that, but also Jackson proposed model checking that could solve the problem of infinite state spaces at the level of software specifications. Abstracting actual states is the method to reduce infinite state spaces.

Despite this sometimes fails to prove a theorem, Jackson claims that this method is sound and it will not prove false theorems. To avoid describing the concrete models which can be huge, Jackson uses a kind of abstract interpretation to approximate the abstract model. This method has been implemented in Standard ML as a prototype checker. If this checker cannot prove a theorem, an abstract counter-example is produced. However, this paper is concerned most in verifying theorems obtained from predicates of a Z specification and how to abstract such theorems. Hardly any discussion is necessary about schemas of such a specification and how to deal with them.

Smith and Winter have proposed an approach that can prove temporal properties of Z specifications systematically [9]. The approach has sequences of processes: transforming the Z specification into an abstract state transition system, verifying that abstract model using a model checker, checking if a false counter-example exists, and refining such an abstract model to be free from such false counter-examples. Properties of certain system are written in LTL theorems and this abstraction works only on one property at each of its processes. This paper describes each of those four steps in detail. It also works on the whole specification, beginning from the state schema until to the last operational schema.

In this paper, as previously stated, the investigation of the abstraction process in model checking Z specification is based on the second paper described above. Those steps are adapted by using the same example given in that paper, but more explanations are given in this paper on each stage of the abstraction process. This paper uses the *Symbolic Analysis Laboratory* (SAL) form of each abstract model to be verified by a model checker chosen. Thus, each abstract deals with its SAL forms. SAL is described more in next section.

3 Verification of Z Specifications

This section provides several sources relates with this research. It begins with Z notation.

3.1 Z Notation

As Z is based on set theory and mathematical logic, this notation can be used to describe formally computing systems. The Z notation is a '*model-based notation*' [10]. A system is modelled by representing its *state*. The scope of the set theory is standard set operators, set comprehensions, Cartesian products, and power sets. Also there are relations, functions, sequences, bags and their operators as well. Moreover, only a first-order predicate calculus is included in the Z notation [11].

To date various tools have been developed and introduced to support Z language, such as **ERZ**, *the CZT*, *Z Word Tools*, **CADiZ**, **fuZZ**, **Z-Eves**, **ZETA**, **Isabelle/HOL-Z**, **ProofPower**, **Vimes**, **z-vimes**, and **Fastest**.

However, there were certain drawbacks, especially in verifying Z specification, associated with those tools were developed from scratch. Martin argues that the issue of tool support is the main obstacle in using Z specification widely [12]. In another

place, Malik et al. argue that with less tools support for Z specifications, only few of them can be used in validating the intended meaning of such Z specifications [13]. Malik et al. have presented the translation of Z to an existing tool, the Alloy Analyser. Not only can the Alloy Analyzer generate instances of Alloy specifications, but also it can model check those specifications. By this translation, it enables 'immediate visual feedback' of such specifications for its designer and this quick response provides a worthwhile verification of Z specifications. Plagge and Leuschel argue that Z is a 'high-level formalism specification' and this might be the cause why Z has limited industrial tools [14]. Derrick et al. argue that to build a model checker directly for a Z specification would take considerable effort due to the abstraction of the language [15]. Jackson argues that the richness of this language might be the issue in verifying it [3].

With these issues in mind, it suggests an alternative method for supporting applications of Z which is to adapt existing tools such as automated checking tools [6]. In most recent studies, there were *ProZ* [14] which has been developed from *ProB*, Bolton [16] who has used the Alloy SAT-solver based counter-example finder to verify data refinements in Z, and Derrick et al. who have been developed Z2SAL [17]. The Z2SAL was intended to use SAL as a model checker and refinement tools for Z specifications which were translated into the SAL input language.

3.2   Model Checking

Model checking does a verification process starting with a model which is described by the user, and discovers whether the hypotheses asserted by the user are valid on the model. This search uses an exhaustive searching of the state space of such a system using suitable graph algorithms [5, 7]. If the model checker cannot satisfy those hypotheses, counter-examples consisting of execution traces will be produced. This automatic generation of counter traces is an important tool in the design and debugging of systems [5].

Temporal logics have proved to be useful for specifying concurrent systems, since they can describe events in ordered time, regardless of time explicitly [5]. The properties used in model checking are based on temporal logic [4] and this was introduced by Clarke and Emerson in their algorithm of temporal logic model checking. In that approach, a formula can be true in some states and false in other states dynamically.

As a model-based approach, the notion of satisfaction is used in model checking. Therefore, the satisfaction test which relates a model and a formula is $M \vdash \varphi$.

According to the particular view of time, temporal logic can be divided into linear time logics and branching time logics. In linear time logics, which are named as *Linear-time Temporal Logic* (LTL), a time is a set of paths, where a path is a sequence of time instances. However, in branching time logics, which are named as *Computation Tree Logic* (CTL), time is represented as a tree, rooted at the present moment and branching out into the future.

3.3 Z2SAL

Derrick et al. [6] said that the idea of translating Z into SAL specifications originated from Smith and Wildman [9], University of Queensland, Australia. However, the current Z2SAL covers broader aspects of Z and tackled optimization issues as well, so it has been extended in a different direction to its *ancestor* [6].

SAL was chosen since it has similar representation to many aspects of Z [15], such as the module mechanism of SAL represents appropriately a Z state transition system [9]. Furthermore, SAL supports expressive mathematics which is necessary for model checking the expressiveness of Z specifications [9]. Moreover, there exists many different tools that use the SAL input language [6] which are offered freely by SRI under academic licence such that attract users to engage in international groups.

Z2SAL translates a Z specification into a SAL module. In this module, it will group a number of definitions including types, constants and modules to describe the states transition system [15]. A SAL module has the general format as follows:

```
State : MODULE =
 BEGIN
  INPUT ...
  LOCAL ...
  OUTPUT ...
  INITIALIZATION [ ... ]
  TRANSITION [
    ...
    ]
END
```

More detail about Z2SAL translation can be found on the related references given above.

3.4 SAL Model Checker

SAL is a framework for combining different tools for abstraction, program analysis, theorem proving and model checking towards the calculation of properties (symbolic analysis) of transition systems [18]. Thus, SAL is used to change the perception and implementation of model checkers and theorem proves which those are based on verification to calculate properties such as abstraction, slicing and composition [19].

As an intermediate language which serves as a medium for representing the state transition semantics of systems with their own source languages, SAL has been integrated with several back-end components. These components are loosely coupled and they relate to each other by using well-defined interfaces [19].

The SAL environment contains a simulator for finite states specifications based on *Binary Decision Diagrams* (BDDs) which allows users to explore different execution paths of a SAL specification [20]. By doing such an exploration, users will be more confident of their model before verification is done on the model.

Regarding model checking, SALenv contains a symbolic model checker called SAL-smc (*simple model checker*). Users can specify properties in LTL and CTL. In

addition to SAL-smc, SALenv also contains SAL-bmc (*bounded model checker*) and it only supports LTL formulas. By using a bounded model checker, SAL can search a state space on a given depth. When a property is invalid, a counter-example will be produced, otherwise, it will be proven.

3.5 Abstraction

It has been shown that a state explosion will limit the applicability of model checking, especially in software verification [5]. The state explosion can involve the usage of memory and time [7].

Previous studies have reported that abstraction is the most important technique for reducing the state explosion problems in model checking [21, 22, 23, 24, 25, 26, 5, 27, 28, and 29]. By using abstraction techniques, a system can be modelled by a smaller system. Thus, if some properties hold for the abstract model, these will hold in the concrete or original model [24].

4 Abstraction on Z Specification

4.1 Introduction

More recently, literature has emerged that Z specifications can be complex enough to be verified. Therefore, abstraction is offered as a possible solution [3, 29]. However, a major problem with that research is the lack of practicality, though some of them are systematic and could be automated.

Abstraction refers to in [29] will group states in the concrete model into equivalence classes, and then map those classes by an abstract function to states in an abstract model. The abstract function has been described in many papers, such as mapping or equivalence relation in [21] and others. However, one that offers systematic method is [25] and this research refers to that method.

In order to be able to prove LTL theorems in a Z specification, such a specification must be restricted to have total operations [29]. This means that every operation must define a post-state, which in some cases is an error state. Details of this abstraction are given in the following section.

4.2 Abstraction Process

This process is adapted from the paper of Smith and Winter [29]. Systematically is described as follows:

Step 1: Abstract generation

To avoid initial value of output by Z2SAL, that output is initialized to 0 for numbered types. To generate an abstract model of a Z specification involves several steps as follows:

1. Define 'monomials' from atomic predicates of properties (LTL theorems). A monomial is a conjunction of, for each predicate, either the predicate or its negation [25]. For example, given one atomic property, $n! \neq v$ (as used in an example on [29]), we have two monomials, which are $n! \neq v$ and $n! = v$.
2. Create state schema whose number of states is equal to the number of monomials.
3. Create abstraction function which is a mapping from monomials to state.
4. Derive initial schema.
5. Derive operational schemas.
6. Derive abstraction version of properties.

These processes need conditions that should hold to show the soundness. Smith and Winter refer to Derrick and Boiten. Simplifications could be done by using a *theorem proves* or with some user guidance in case of more complicated specification [29].

Step 2: False counter-example detection
If the property can be proved or the counter-example belongs to the original specification, then go to Step 4. If this is not the case, refinement takes place. This brings us to Step 3. The approach that is used to detect a false counter-example is based on [28], as follows:

1. Find a sequence of operations in the concrete specifications that passes through concrete states related to the abstract states in the counter-example.
2. If it cannot be found, then the corresponding states are not reachable in the concrete model, and the counter-example is indeed a false counter-example. This is indicated by **False** value in the simplified predicate of one of those schemas.

Step 3: Refine the abstract model
This step involves a number of refinements that are called if the refined property cannot be proved. This is needed to avoid the false counter-example during processing with the model checker. Smith and Winter uses approach from [28]:

1. Split the abstract state that is related to the last reachable concrete state (using abstraction function) into a state that has transition to the next state, and another state that has no such transition. Presuming the last reachable state is t_i and the next state is t_{i+1} in the abstract model. The approach that was used by Smith and Winter is if the last reachable concrete state is schema C_i, then schema C_{i+1} is the schema that has false in its predicate. Thus, state t_i will be replaced by two values, which are states s_{n+1} and s_{n+2}, and the abstraction function will also be modified. The predicate $p \Rightarrow s = t_i$ will be replaced by two new predicates:

$$p \wedge (\exists \; CState' \cdot (COp\_1 \vee ... \vee COp_n) \wedge q \Rightarrow s = s_{n+1} \tag{1}$$

and

$$p \wedge (\neg \; \exists \; CState' \cdot (COp\_1 \vee ... \vee COp_n) \wedge q \Rightarrow s = s_{n+2} \tag{2}$$

2. Refined the abstract specification and abstract properties based on those new abstraction function
3. Check whether or not there still exists a false counter-example. If this the case, then do as described in Step 2. Otherwise, terminate the process as the real counter-example has been found or the properties could be proved.

In this paper, an investigation has been carried out that begun by providing a Z specification, translating it into a SAL input file, adding an LTL theorem, and conducting the abstraction process. The next section will provide an experiment on that process.

4.3 Experiment

The following example is taken from pages 268-269 of [29]. However, the abstraction process is given in detail until the last step where there is no need to detect false counter-examples.

Providing a Z Specification. This system deals with allocating a positive unique number from a requester and sending it back to them. A Z specification below represents that system.

$$
\begin{array}{|l}
\hline
\textit{Allocator} \\
\hline
\textit{used}: \mathbb{P}\mathbb{N}_1 \\
\textit{alloc}: \;\cdot\; \mathbb{N}_1 \\
\hline
\#\textit{alloc} \leq 1 \\
\hline
\end{array}
$$

$$
\begin{array}{|l}
\hline
\textit{Init} \\
\hline
\textit{Allocator} \\
\hline
\textit{used} = \varnothing \\
\textit{alloc} = \varnothing \\
\hline
\end{array}
$$

Request

Δ Allocator

$(alloc = \varnothing \wedge used \neq \mathbb{N}_1) \Rightarrow (\exists\ n: \mathbb{N}_1\ \cdot\ n \notin used\ \wedge$

$alloc'\quad = \{n\} \wedge used = used\ \cup\ \{n\})$

$(alloc \neq \varnothing \vee used = \mathbb{N}_1) \Rightarrow (alloc'\quad = alloc \wedge used'\quad = used)$

Send

Δ Allocator

$n!: \mathbb{N}$

$alloc = \{n!\} \Rightarrow alloc'\quad = \varnothing \wedge used'\quad = used$

$alloc = \varnothing \Rightarrow n! = 0 \wedge alloc'\quad = alloc \wedge used'\quad = used$

Verifying by SAL Model Checker. In this step, there are several processes, starting by translating the specification into a SAL file, adding an LTL theorem, and verifying by SAL model checker. The translation is done by Z2SAL.

Translation into SAL is to enable the specification to be verified by SAL model checker. Then, add some properties in an LTL theorem form to that SAL file. This step is just for the sake of confidence in the system. However, this SAL is not given here due to the limitation on number of pages. Once the system functions properly, the next step is taken. Otherwise, such a system must be revised first.

The translation of the Z specification into a SAL file is the main difference in this paper from work done by Smith and Winter. Thus, in this paper, verification is conducted on the SAL file rather than directly on the Z specification.

For that specification, a property got from that paper said that '*it is never the case that the same number will be sent more than once*', and this should be proved:

$$\vdash G(n\_ \neq v \vee X(G(n\_ \neq v))) \tag{3}$$

Two is chosen as an instance of v since v is a non-zero natural number. Jackson said that the counter-example mostly will be found at types whose number of instances is three [30].

The next process is to verify it with SAL model checker. However, a counter-example was produced that indicates either this system is not correct or the translation is not correct.

Based on that counter-example, the translator initialized output variable to a value equals to the maximum size of that set which is 2, whereas in [29] every numbered output variable is initialized to $\perp$ as well as the post-state of that variable in operational schemas which does not declare any output or input. As for that example, it has been done in *Request* schema. Thus, initial value for such an output variable

was set to 0 and in the *Request* transition it was also defined to be 0 its post-operational value.

Those changes were written manually and directly on the generated SAL file. This revised SAL was re-run by the SAL model checker, and a counter-example was generated. This time, a number that has ever been sent can be sent again since the *Send* transition deletes such a number from its set. Thus, the *Send* schema was modified as follows:

Send

Δ *allocator*

$n!: \mathbb{N}$

$alloc \neq \varnothing \Rightarrow n! \in alloc \land alloc' = \varnothing \land used' = used$

$alloc = \varnothing \Rightarrow n! = 0 \land alloc' = alloc \land used' = used$

However, another counter-example was generated. It seems that there is an anomaly in the SAL model checker as the *ELSE* part was fired, although it should have been not, and this part was followed by other transitions. It is supposed to be the *ELSE* part, if it were exists, occurs in the end. Thus, the *ELSE* part was deleted, and finally the SAL file was proved.

Abstraction Process. To get more advantages from abstraction, the abstract specification must be simplified first. The simplified version of that abstract model has been given in [29]. The SAL file generated by Z2SAL of that abstract model is given below:

```
AllocatorAbs : CONTEXT = BEGIN
STATE : TYPE = DATATYPE
 s1, s2
END;
State : MODULE = BEGIN
 LOCAL s : STATE
 INITIALIZATION [ s = s1 --> ]
 TRANSITION [  Request : s' = s1 --> s' IN {x : STATE | TRUE}
  [] Send : TRUE --> s' IN {x : STATE | TRUE}
  [] ELSE --> ]
 END;
 th1:THEOREM State |- G(s=s1 OR X(G(s = s1)));
END
```

As given on that paper, this abstract model will produce a counter-example on that property, which is not supposed to be occurred. The sequence of that counter-example is ⟨ s1, s2, s2, s1 ⟩.

First Detection and Refinement. The next step is to detect whether or not such a counter-example is a false counter-example. Indicated by *False* in the predicate of the

C1 schema, that counter-example is a false one [29]. The state *s1* will be split into *s3* and *s4* states.

The first refinement on that abstract model has been done. The resultant SAL file, which is modified from the previous abstraction SAL, is given as follows:

```
...
STATE : TYPE = DATATYPE
 s3, s4, s2
...
 INITIALIZATION [s = s4
...
  Request :   (s=s3 => s'=s3) AND ((s=s2 OR s=s4) => (s'=s3 OR s'=s4))
...
  Send :   (s=s3 => s'=s2) AND (s=s4 => s'=s4)
...
 th1:THEOREM State |- G((s=s3 OR s=s4) OR X(G((s=s3 OR s=s4))));
END
```

In the SAL above, some information is hidden since it is the same as the previous SAL. A counter-example was produced with sequence・s4, s3, s2, s2, s2・.

Second Detection and Refinement. That counter-example will be detected to prove whether it is a false counter-example or not. This process is represented by these schemas:

┌─ *C0* ─────────────────────────────────
│ *InitIO*
│ ────────────────────────────────
│ $\exists$ *s*: *State* ・
│   (*s* = *s4*)
│   ($n!$ $\neq$ 2 $\wedge$ *alloc′*   = {2} $\Rightarrow$ *s* = *s3*)
│   ($n!$ $\neq$ 2 $\wedge$ *alloc′* $\neq$ {2} $\Rightarrow$ *s* = *s4*)
│   ($n!$ = 2 $\Rightarrow$ *s* = *s2*)
└───────────────────────────────────────

This schema will be simplified into:

┌─ *C0* ─────────────────────────────────
│ *InitIO*
└───────────────────────────────────────

since *n!* $\neq$ 2 and *alloc′* $\neq$ {2} in *Init*.

The *C1* is simplified into *TRUE* since *s=s3*, *n!* $\neq$ 2 and *alloc′* = {2} are *TRUE*. There is a post-state of *Request* operation from pre-state *C0* to this schema.

The *C2* schema was simplified into *TRUE* on its predicate. Since *s=s2* and *n!* = 2, *Send* operation is selected. It tells us that there is a post-state of *Send}* from pre-state *C0* via *C1* state and it ends up in *C2* state.

The *C3* schema simplified into *FALSE*, this means that this concrete schema cannot be reached from the initial schema via concrete schemas both *C1* and *C2*. Since *n!* = 2 and *alloc* = ∅ from *C2*, none of those operation could be selected. Thus, the last reachable concrete schema is *C2* with *s2* state and it will be split into two states which are *s5* and *s6*. As given on [29], *s5* is unreachable since there is no transition either from *s4* or *s3* or from both of them to *s5*.

The second refinement results this SAL file:

```
...
STATE : TYPE = DATATYPE
 s4, s6, s3
...
 INITIALIZATION [
   s = s4
...
  Request :   (s=s3 => s'=s3) AND ((s=s4 OR s=s6) => (s'=s3 OR s'=s4))
...
  Send :    (s=s3 => s'=s6) AND ((s=s4 OR s=s6) => s'=s4)
...
 th1:THEOREM State |- G((s=s3 OR s=s4) OR X(G((s=s3 OR s=s4))));
END
```

The sequence of states on its counter-example is ⸳ s4, s3, s6, s6, s3 ⸳ .

Third Detection and Refinement. The *C0* schema will be simplified into *TRUE* shown by this schema:

C0 _____

InitIO
_____

$\exists s$: *State* ⸳

  $s = s4$

  $n! \neq 2 \wedge alloc' \ = \{2\} \Rightarrow s = s3$

  $n! \neq 2 \wedge alloc' \ \neq \ \{2\} \Rightarrow s = s4$

  $n! = 2 \wedge alloc' \ \neq \{2\} \Rightarrow s = s6$

since *n!* ≠ 2 and *alloc'* ≠ {2} in *Init*.

The *C1* is simplified into *TRUE* since *s = s3*, *n!* ≠ 2 and *alloc'* = {2}. The process of detection results a *TRUE* in *C2* schema. In this schema, *s=s6*, *n!* = 2 and *alloc'*≠ {2}. From *C1* it gets *alloc* = {2} and *n!'* = 0. This will fire *Send* operation.

C3 schema will result *FALSE* since it gets *n!'* = 2 and *alloc'* = ∅ from *C2* and so it cannot select any of those operation. Thus, the s6 will be split into *s7* and *\s8* states. The latter is the unreachable state.

The third refinement produces this SAL:

```
...
STATE : TYPE = DATATYPE
```

```
 s4, s7, s3
...
 INITIALIZATION [
    s = s4
...
  Request :
   (s=s3 => s'=s3) AND (s=s4 => (s'=s3 OR s'=s4)) AND (s=s7 => s'=s4)
...
  Send :   (s = s3 => s' = s7) AND ((s = s4 OR s=s7) => s' = s4)
...
 th1:THEOREM State |- G((s=s3 OR s=s4) OR X(G((s=s3 OR s=s4)))));
END
```

The sequence of states in the generated counter-example is ・ s4, s3, s7, s7, s3 ・.

Fourth Detection and Refinement. The detection got *TRUE* from *C0* to *C3* and *FALSE* in the *C4* schema. It requires the *s4* state to be divided into *s9* and *s10* states.

C0 before simplification is showed as follows:

C0

InitIO

$\exists\, s: STATE\ ・$

$s = s4$

$n! \neq 2 \wedge alloc' \quad = \{2\} \Rightarrow s = s3$

$n! \neq 2 \wedge alloc' \quad \neq \{2\} \Rightarrow s = s4$

$n! = 2 \wedge alloc \neq \{2\} \wedge 2 \in used \Rightarrow s = s7$

The fourth refinement results this following SAL:

```
...
STATE : TYPE = DATATYPE
 s9, s10, s7, s3
...
 INITIALIZATION [
    s = s9}
...
  Request :
   (((s=s3 OR s=s9) => s'=s3) AND ((s=s7 OR s=s10) => (s'=s10))
...
  Send :
   (s=s3 => s'=s7) AND (s=s9 => (s'=s9) AND ((s=s7 OR s=s10) => (s'=s10))
...
 th1:THEOREM State |- G((s=s3 OR s=s9 OR s=s10) OR X(G((s=s3 OR s=s9 OR
s=s10)))));
END
```

Finally, this SAL was proved by SAL model checker. This means that such a property is a true behavior of that system.

```
$ sal-smc AllocatorAbs4 th1
WARNING: SALPATH is now called SALCONTEXTPATH.
proved.
```

4.4 Results and Discussion

Based on the experiment, the abstraction process given on the reference can be followed in this paper. The result can also be reproduced appropriately. Nevertheless, this paper brings insight into the practicality of integration of several tools in model checking Z specification. Thus, this paper gives a different view of such an abstract model which is to view it as a SAL file not as a Z specification.

5 Future Works and Conclusion

This paper has given a practical demonstrator of the abstraction process on a Z specification though some of it is manual. Although this work has been carried out on the example given on [29], several adjustments have been made into the original specification. This paper has utilized the Z2SAL translator and SAL model checker as well.

After this step, the future work will be the investigation of which parts of such an abstraction process could be automated either as an extension of Z2SAL or as a different tool outside Z2SAL. Furthermore, this investigation will be implemented on several new examples.

To conclude, the abstraction process should offer significant value for verifying systems, specifically those which have large states spaces finite systems or even the infinite ones. This method will also enrich the available tools on verification of Z specification.

References

1. Potter, B., Till, D., and Sinclair, J.: An introduction to formal specification and Z. Prentice Hall PTR (1996)
2. West, M.M.: Issues in Validation and Executability of Formal Specifications in the Z Notation. Thesis of University of Leeds (2002)
3. Jackson, D.: Abstract model checking of infinite specifications. In: Naftalin, M., Denvir, BT, and Bertran, M. (eds.) FME'94: Industrial Benefit of Formal Methods. LNCS, vol. 873, pp. 519—531. Springer, Heidelberg (1994)

344

4. Huth, M. and Ryan, M.: Logic in Computer Science: Modelling and reasoning about systems. Cambridge University Press, (2004)
5. Clarke, E.M., Grumberg, O. and Peled, D.A.: Model checking. MIT Press (1999)
6. Derrick, J., North, S., and Simons, A.J.H.: Z2SAL: a translation-based model checker for Z. Formal aspects of computing. 23 1, 43—71. Springer (2011)
7. Pelánek, R.: Reduction and Abstraction Techniques for Model Checking. PhD thesis, Faculty of Informatics, Masaryk University, Brno (2006)
8. Merz, S.: Model checking: A tutorial overview. Modeling and verification of parallel processes. LNCS, vol. 2067, pp. 3—38. Springer, Heidelberg (2001)
9. Smith, G. and Wildman, L.: Model checking Z specifications using SAL. In: Treharne, H., King, S., Henson, MC, Schneider, SA (eds.) ZB 2005: Formal Specification and Development in Z and B. LNCS, vol. 3455, pp. 85—103. Springer, Heidelberg (2005)
10. Jacky, J.: The way of Z: practical programming with formal methods. Cambridge University Press, (1996)
11. Woodcock, J. and Davies, J.: Using Z: specification, refinement, and proof. Prentice-Hall, Inc. (1996)
12. Martin, AP: Proposal: Community Z Tools Project (CZT). Accessed from http://www.cs.ox.ac.uk/people/andrew.martin/CZT/proposal.html (2001)
13. Malik, P., Groves, L. and Lenihan, C.: Translating z to alloy. ASM, Alloy, b and Z. In: Frappier, M., Glasser, U., Khurshid, S., Laleau, R., and Reeves, S. (eds.) LNCS, vol. 5977, pp. 377—390. Springer (2010)
14. Plagge, D. and Leuschel, M.: Validating Z specifications using the ProB animator and model checker. In: Davies, J., and Gibbons, J. (eds.) Integrated Formal Methods. LNCS, vol. 4591, pp. 480—500. Springer (2007)
15. Derrick, J., North, S., and Simons, A.J.H.: Issues in implementing a model checker for Z. In: Zhimming Liu and Jifeng He (eds.) Formal Methods and Software Engineering. LNCS, vol. 5238, pp. 678—696. Springer (2006)
16. Bolton, C.: Using the alloy analyzer to verify data refinement in Z. Electronic Notes in Theoretical Computer Science. Elsevier, 137 2, 23--44 (2005)
17. Derrick, J., North, S. and Simons, A.J.H: Z2SAL-building a model checker for Z. In: Borger, E., Butler, M., Bowen, JP, and Boca, P. (eds.) Abstract State Machines, B and Z. LNCS, vol. 5238, pp. 280—293. Springer (2008)
18. De Moura, L., Owre, S. and Shankar, N.: The SAL language manual. Computer Science Laboratory, SRI International, Menlo Park, CA, Tech. Rep. CSL-01-01 (2003)
19. Bensalem, S., Lakhnech, Y. and Owre, S.: Computing abstractions of infinite state systems compositionally and automatically. In: Alan JH, Vardi, MY (eds.) Computer Aided Verification. LNCS, vol. 1427, pp. 319—331. Springer (1998)
20. de Moura, L.: SAL: tutorial. Computer science laboratory, SRI International (2004)
21. Clarke, E.M., Grumberg, O. and Long, D.E.: Model checking and abstraction. ACM Transactions on Programming Languages and Systems (TOPLAS). 16 5, 1512--1542 (1994)
22. MVller, O. and Nipkow, T.: Combining model checking and deduction for I/O-automata. Tools and Algorithms for the Construction and Analysis of Systems. Springer, 1--16 (1995)
23. Havelund, K. and Shankar, N.: Experiments in theorem proving and model checking for protocol verification. In: Gaudel, MC, Woodcook, J. (eds.) FME'96: Industrial Benefit and Advances in Formal Methods. LNCS, vol. 1051, pp. 662—681. Springer (1996)
24. Dams, D., Gerth, R., and Grumberg, O.: Abstract interpretation of reactive systems. ACM Transactions on Programming Languages and Systems (TOPLAS) 19 2, 253--291 (1997)
25. Graf, S. and Saïdi, H.: Construction of abstract state graphs with PVS. In: Grumberg, O. (eds.) Computer aided verification. LNCS, vol. 1254, pp. 72—83. Springer (1997)

26. Daws, C. and Tripakis, S.: Model checking of real-time reachability properties using abstractions. In: Steffen, B. (eds.) Tools and Algorithms for the Construction and Analysis of Systems. LNCS, vol. 1384, pp. 313—329. Springer (1998)
27. Graf, S.: Characterization of a sequentially consistent memory and verification of a cache memory by abstraction. Distributed Computing. Springer, 12 2-3, 75--90 (1999)
28. Clarke, E., Grumberg, O., Jha, S., Lu, Y. and Veith, H.: Counterexample-guided abstraction refinement. In: Emerson, EA, and Sistla, AP (eds.) Computer aided verification. LNCS, vol. 1855, pp. 154—169. Springer (2000)
29. Smith, G., and Winter, K.: Proving temporal properties of Z specifications using abstraction. In: Bert, D., Bowen, JP, King, S., and Walden, M. (eds.) ZB 2003: Formal Specification and Development in Z and B, Springer 260--279, (2003)
30. Jackson, D.: Alloy: a lightweight object modelling notation. ACM Transactions on Software Engineering and Methodology (TOSEM). 11 2, 256--290 (2002)

Lipschitz cover based image colorization

Adam Popowicz[1] , Boguslaw Cyganek[2], Bogdan Smolka[1]

[1] Silesian University of Technology, Department of Automatic Control, Akademicka 16,
44-100, Gliwice, Poland, Adam.Popowicz@polsl.pl, Bogdan.Smolka@polsl.pl
[2] AGH University of Science and Technology, Al. Mickiewicza 30,
30-059 Krakow, Poland, cyganek@agh.edu.pl

Abstract. In the paper we present a novel method of grayscale image colorization utilizing the idea of Lipschitz cover applied to a digital image. Our algorithm represents the interactive type of colorization in which the user indicates the hints in form of scribbles of a given color. The method colorizes the grayscale image by analyzing the intensity distance from the scribbles using the Lipschitz upper cover. Comparing with other methods, our proposition is much faster, simpler and it automatically obtains the weights of colors from the inserted scribbles needed for the final color estimation.

Keywords: image colorization, Lipschitz cover, image processing.

1 Introduction

Image colorization is a process of introducing colors to a grayscale image. In the case of one-dimensional pixel value - brightness - the colorization algorithms introduce three values. Depending on the color model, they may be represented by e.g. R, G and B channels (Red, Green and Blue) or Y, Cr, Cb channels (luminance, red chrominance and blue chrominance).

A colorized image is obviously more attractive for a viewer. Additionally it contains more details which are easier noticeable. Colorization is crucial in some important applications like in the medicine where the colorized versions of the images carry more information and bring a lot of educational benefits [13]. For imaging in non-visible light ranges colorization may provide better visualization of the analyzed scene, like in the infrared imaging [10].

The article is organized as follows. In the next section we enumerate and discuss the most important methods used for the colorization. In Section 3 we present the principle of our algorithm and present its details. In the next section we show some examples of resulted colorized images in several applications and compare the results of our colorization with competing algorithms. Finally we draw some conclusions in Section 6.

2 Brief overview of modern colorization algorithms

There is no direct solution for the transformation from the one-dimensional space (brightness only) to three dimensions (brightness and chrominance), thus the colorization is not a trivial task and can be performed in several ways. Two most important are the automatic and the semi-automatic approach.

In the first one, the user has to provide the reference color image. The algorithm, by proper color matching, transfers the colors from the reference image to the grayscale one. Although there is no need of interaction, such methods do not provide the solution for enhancing or modifying the colorization final result. They also require properly chosen source image, which is impractical when e.g. the ultrasound image is to be colorized. Some examples of the algorithms obtaining high colorization quality with the use of automatic approach are described in [10, 13, 14].

The second method - semiautomatic - requires the indication of the color hints called scribbles. The user has to draw the scribbles of required color in particular part of the image. The colorization process assigns to the pixels the colors of the closest scribble. Both, the metric and/or the brightness distance is usually considered. Our method of colorization is of the semi-automatic type thus we present more details of some competing methods.

In [6] the authors present the utilization of optimization tools for colorization purpose. They assume that neighboring pixels of similar intensity should also have similar color. The proposed cost function may be minimized using standard optimization techniques. The algorithm was implemented in Matlab environment and made available online thus it was used for comparison in several similar publications.

The article [11] shows the method of color blending with initial calculations of the distances from scribbles using the Dijkstra algorithm. The solution was proved to be faster than the one presented in [6] and provides very high colorization quality. However, there is no available code for this method, thus we had to implement it to compare the results of this colorization with the outcomes of our algorithm.

Another colorization method was presented in [5]. The authors utilize the distance transformation which allows for fast calculations of the distance from the scribbles inserted by the user. They consider the superposition of spatial distance and intensity difference to calculate the hybrid distance maps for each scribble. By proper color blending using the weights from aforementioned distances they show very realistic colorization results. This method was also implemented for use in the experimental part of our paper.

Several other methods, not compared with our algorithm, are given in the literature. The utilization of fuzzy clustering can be found in [12]. In [2] the authors use the oversegmentation of grayscale images and they gradually join the segments as long as they achieve the number of segment equal to the number of the scribbles. The textural features of the image are also exploited in [3, 4].

3 Lipschitz cover

The utilization of Lipschitz cover for shading-off effect was suggested in [7]. The cover is the equivalent to the opening operation on the grayscale image with a cone structuring element [1]. The authors of [9] present a fast method of lower and upper Lipschitz cover calculation on a digital image. The idea is based on the distance transformation and the double scan algorithm which is computationally simple and efficient. We adopted this idea in our concept of the colorization.

The main part of the algorithm is based on the double scan distance calculation algorithm [8], which enables quick distance estimation from binary objects of a given image. It performs the running analysis of five pixels in local moving window which moves through the image in two scans. In the first one the direction is from the left to right, up to bottom. Then in second scan it moves in opposite direction - from the right to left, bottom to top. The mask depends on the scanning direction of shifting and is presented in Fig. 1.
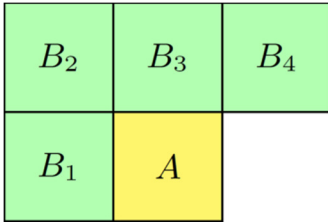
(a) Mask utilized in the first scan.   (b) Mask utilized in the second scan.
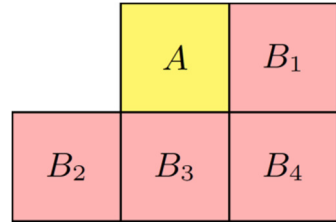
Fig. 1. Masks utilized in double-scan algorithm.

The main part of The distance map is first initialized with zeros at pixels forming a binary object while the others are assigned infinity as depicted in Fig. 2. Then the mask runs in two scans performing simple calculations in each step as defined in following equation:

$$D_{new}(A) = min\{D(B_1), D(B_2), D(B_3), D(B_4), D(A)\}, \tag{1}$$
$$D(A) = D_{new}(A)$$

where $D(A)$ is the distance value for pixel A, min is the minimization process, $D_{new}(A)$ is new estimation of distance of pixel A which replaces the previous value $D(A)$.

We present the simplified version of its mathematical representation so that it is better understandable for the reader. The results of performing the double scan algorithm on the digital image with indicated binary object is presented in Fig. 2.
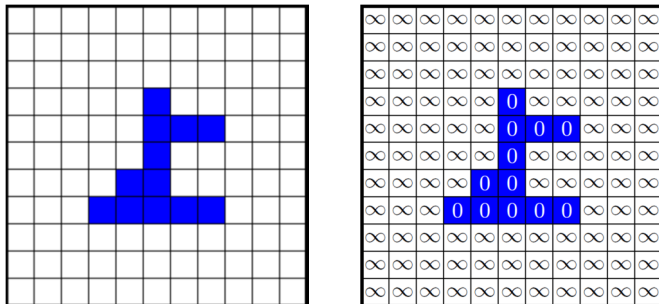
(a) Image with indicated feature. (b) Distance map initialization.
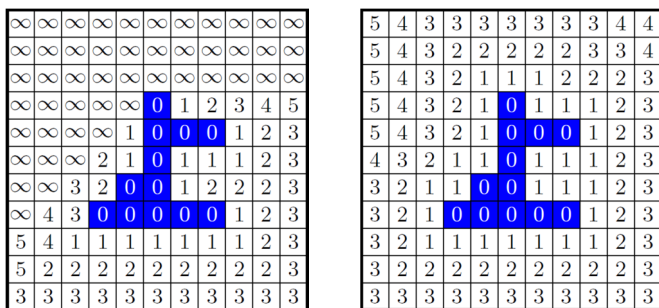
(c) Distance map after 1st scan. (d) Distance map after 2nd scan.

Fig. 2. Example of double-scan algorithm.

The similar procedure was utilized in [9] for the calculation of lower and upper Lipschitz cover calculations. For this purpose the authors use the intensity value of each pixel. When initializing the distance matrix, each pixel containing the binary image is given the distance equal to its intensity. The others are given 0 when calculating upper cover or infinity for lower cover. Additionally the formula (1) is modified depending on the cover: for the upper – (2), for the lower – (3).

$$D(A) = max\{(D(B_i) - s), D(A) \text{ for } i = 1, 2, 3, 4\}, \tag{2}$$

$$D(A) = min\{(D(B_i) + s), D(A) \text{ for } i = 1, 2, 3, 4\}, \tag{3}$$

where s is the so-called *slope* constant.

In our colorization algorithm we decided to follow the steps of upper Lipschitz cover. In colorization case the feature is the scribble indicated by a user so that all pixels within a scribble are given distances equal to their grayscale intensities. Others are set to zero. We modified the slope value so that it represents the intensity difference between each pixel B_i and A (4) and included two parameters: and which may be modified to obtain better colorization results. They represent the weight of metric ( ) or intensity ( ) distance in our calculations:

$$s_i = \left(\frac{I(A) - I(B_i)}{\lambda}\right)^2 + \delta_i \tag{4}$$

$$D(A) = max\{(D(B_i) - s_i), D(A) \text{ for } i = 1, 2, 3, 4\},$$

The final value of a color pixel is calculated using simple weighting average, where weights are taken from the aforementioned distances from a scribble. The exemplary YCrCb representation of a pixel color is given as:

$$Cr = \frac{\sum_{i=1}^{N} Cr_i \, D_i}{\sum_{i=1}^{N} D_i}, \qquad Cb = \frac{\sum_{i=1}^{N} Cb_i \, D_i}{\sum_{i=1}^{N} D_i}, \qquad Y_{color} = Y_{grayscale}, \tag{5}$$

where Cr_i, Cb_i is the chrominance representation of the i-th scribble indicated by a user, D_i is the distance from the the i-th scribble, Y is the luminance value (in this colorization process it is transferred from the grayscale to the color image).

4 Colorization optimization

To achieve the best colorization results, the user is given two parameters which can be adjusted. They are: , which is responsible for increasing the intensity impact on distance calculations and   , which is the step penalty. They both should be carefully tuned to optimize the colorization effects.

To show the impact of parameter values on the colorization quality, we show an exemplary color image (Fig. 3a) which was converted to grayscale and then, after scribbling (Fig. 3b), it was colorized using our method. To estimate the colorization quality we used PSNR (Peak Signal to Noise Ration) calculated using original image as a reference. We investigated PSNR for R, G and B channels individually (PSNR$_R$, PSNR$_G$, PSNR$_B$) and we also provided the mean value PSNR$_{mean}$.

$$PSNR_k = 10 log_{10}\left(\frac{255 \cdot 255}{\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}[p_r^k(i,j) - p_c^k(i,j)]^2}\right) \tag{6}$$

$$PSNR_{rgb\ mean} = \frac{PSNR_R + PSNR_G + PSNR_B}{3}, \tag{7}$$

where k – channel indicator (e.g. R, G or B in RGB color model), N, M – width and height of the image, $p_r^R(i,j)$, $p_r^G(i,j)$, $p_r^B(i,j)$ – the intensity of pixel at position (i, j) in reference image (channel R, G, B), $p_r^R(i,j)$, $p_r^G(i,j)$, $p_r^B(i,j)$ – the intensity of pixel at position (i, j) in colorized image.

The plots presented in Fig. 4 allow the user to find proper proportion between   and   . Exemplary colorized images for proper and wrong parameter selection are shown in Fig. 5. In Fig. 5a we present the effect of a too low value of   in respect to   , which leads to many not colorized pixels visible in form of gray fields. In comparison, we show in Fig. 5b the poor outcome of the colorization with too large   , which results in very strong color blending. Finally, in Fig. 5c we present the colorized image using optimal parameter selection (= 2.9,   = 12).

(a) Original image
(250 x 250 pixels)

(b) Grayscale version
scribbled with 7 colors
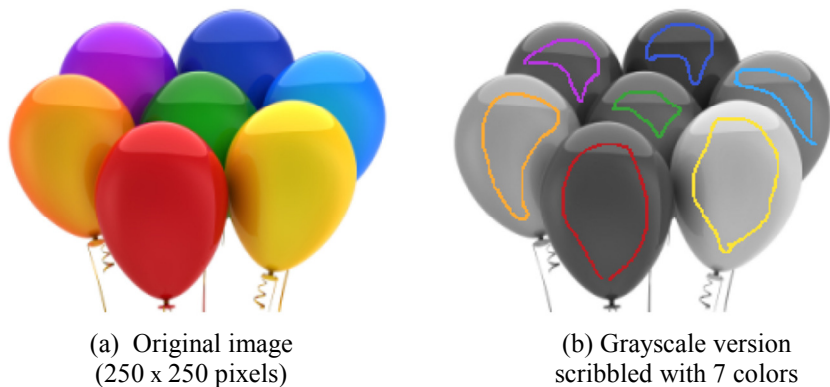
Fig. 3. Original and scribbled image for parameter optimization tests.
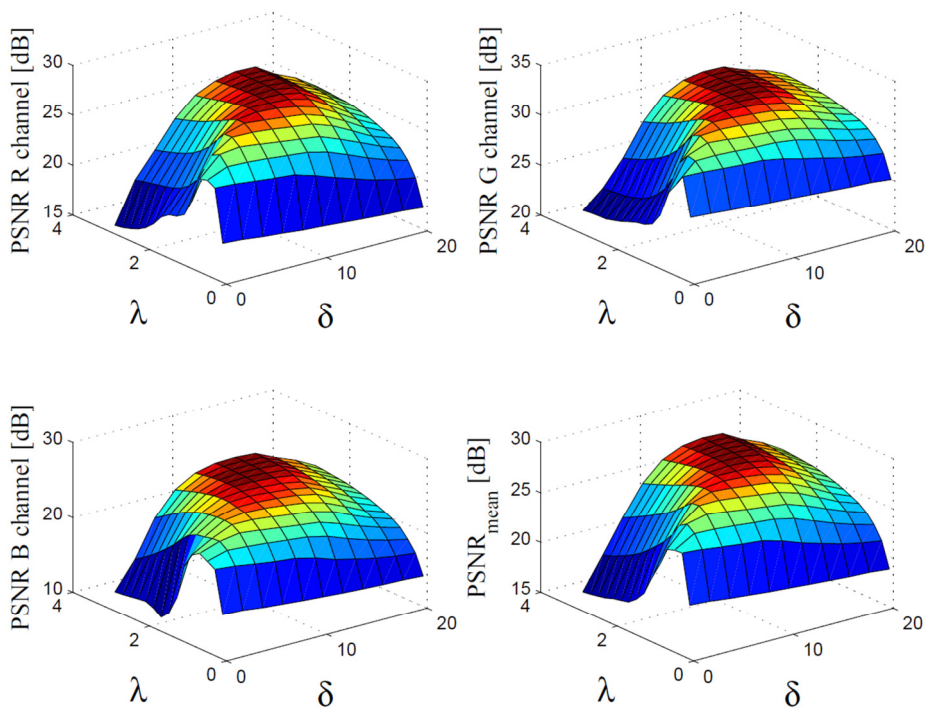
Fig. 4. Dependence of PSNR on λ and δ for the test image depicted in Fig. 3.

(a) Too large intensity impact (= 0.1, = 0). (b) Too small intensity impact (= 2.5, = 0). (c) Proper parameter selection (= 2.9, = 12)
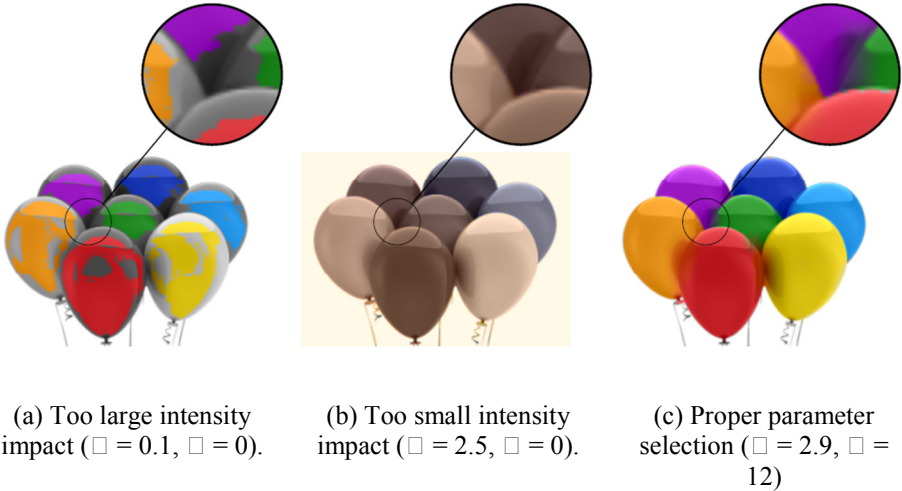
Fig. 5. Exemplary image colorization for different parameter selection.

5 Comparison with other methods

We compared the result of colorization using our method with colorization algorithms found in the literature. For the first algorithm – *Colorization Using Optimization* [6] – we used Matlab code available in on-line resources. The other chosen methods are: *Chrominance Blending* [11] and the colorization utilizing distance transformations [5]. They were implemented based on their pseudo-codes in corresponding publications.

The colorization process of each method was optimized by tuning available parameters to achieve best outcomes. We have run the procedures on a PC (Intel Core i5, 16GB RAM) to check calculation time for several images with different number of indicated scribbles. The results are presented using images exhibited in Fig. 6. The calculation times are given in Table 1.

As one can see, each algorithm is able to provide very realistic colorization results. Some minor differences, which in general do not have much impact on the overall quality, are pointed by a zoom tool. However, the calculations times for the methods differ significantly. We found that our algorithm works with similar, high speed as the one utilizing distance transformations [5]. Two remaining methods show higher computational burden. While *Colorization Using Optimization* [6] seems to be not dependent on the number of indicated colors/scribbles, *Chrominance Blending* [11] is extremely sensitive to this variable. However, it is to be added that *Chrominance Blending* algorithm was implemented in its basic version and it is probably the reason for such low efficiency. It is significant that our algorithm demonstrated mostly the lowest computational times among analyzed methods. It is an important factor as the colorization process is usually repeated iteratively by a human operator to enhance the result by better scribbling.
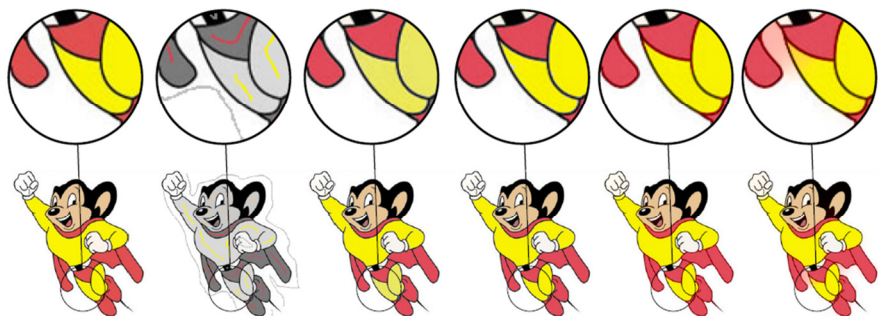
(a) 300x291. (b) 4 colors. (c) Proposed. (d) [5]. (e) [11]. (f) [6].
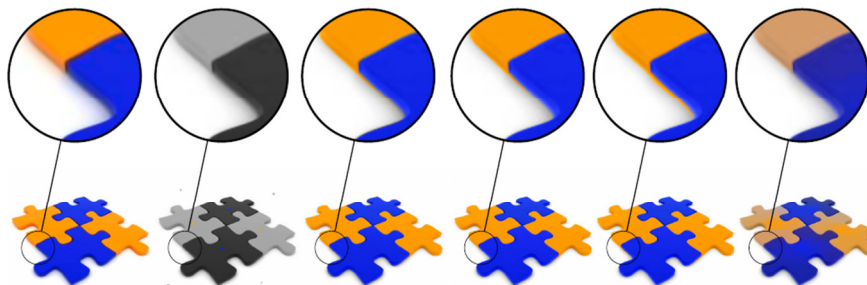
(g) 300x291. (h) 4 colors. (i) Proposed. (j) [5]. (k) [11]. (l) [6].
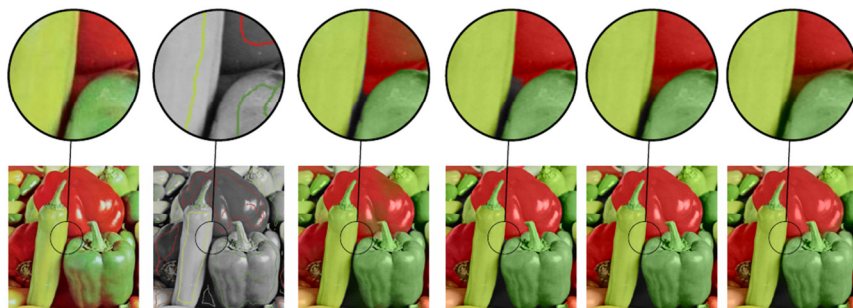
(m) 300x291. (n) 4 colors. (o) Proposed. (p) [5]. (q) [11]. (r) [6].

Fig. 6. Comparison results: (a, g, m) original image with provided resolution, (b, h, n) scribbled image with number of colors, (c, i, o) colorization using our algorithm, (d, j, p) colorization using *Distance Transformation* [5], (e, k, q) *Chrominance Blending* [11], (f, l, r) *Colorization Using Optimization* [6].

354

Table 1. Calculation time for colorization of images depicted in Fig. 6.

	Image 1	Image 2	Image 3
Image resolution	300 x 291	460 x 460	480 x 480
Scribbled colors number	4	3	25
Proposed method	2.2 s	3 s	22.3 s
Distance transformation [5]	2.3 s	3 s	23.8 s
Chrominance Blending [11]	14.5 s	9.2 s	360 s
Colorization Using Optimization [6]	5 s	15 s	14.7 s

6 Applications

The colorization algorithm can be utilized in various applications. In this article we show some examples of using our novel method (see Fig. 7). The applications include: astronomy, biology, medicine, night-vision, old photos and old grayscale movie colorization. While for some of them the colorization plays only minor or additional role, for others - like in the medicine - it may be of a special interest.

7 Conclusions

In the paper we have presented a novel image processing algorithm enabling the colorization of grayscale images. The method utilizes the concept of upper Lipschitz cover combined with a fast distance transformation and double-scan algorithm.

We compared our results of colorized images with the outcomes of competitive, well-established interactive colorization algorithms. The method proved its high colorization quality and computational efficiency.

In the paper we have also presented several interesting applications of our method. They include astronomy, medicine, biology and others. The use of the algorithm in segmentation process is worth of consideration and will be investigated in the future.

Acknowledgment

| Original image. | Scribbles composition. | Colorized image. |

Fig. 7. Examples of image colorization using the proposed method.

References

1. H. Breu, J. Gil, D. Kirkpatrick, and M. Werman. Linear time euclidean distance transform algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(5):529-533, May 1995

2. V.G. Jacob and S. Gupta. Colorization of grayscale images and videos using a semiautomatic approach. In 2009 16th IEEE International Conference on Image Processing (ICIP), pages 1653-1656, Nov 2009.

3. M. Kawulok, J. Kawulok, and B. Smolka. Discriminative Textural Features for Image and Video Colorization. IEICE Transactions on Information and Systems, 95:1722-1730, 2012.

4. M. Kawulok and B. Smolka. Texture-adaptive image colorization framework.EURASIP J. Adv. Sig. Proc., 2011:99, 2011.

5. P. Lagodzinski and B. Smolka. Application of the extended distance transformation in digital image colorization. Multimedia Tools Appl., 69(1):111-137, March 2014.

6. A. Levin, D. Lischinski, and Y. Weiss. Colorization using optimization. ACM Trans. Graph., 23(3):689-694, August 2004.

7. P. Moreau and Ch. Ronse. Generation of shading-o_ in images by extrapolation of Lipschitz functions. Graphical Models and Image Processing, 58(4):314-333, 1996.

8. F.Y.-C. Shih and O. R. Mitchell. A mathematical morphology approach to euclidean distance transformation. IEEE Transactions on Image Processing, 1(2):197-204, Apr 1992.

9. M. Stencel and J Janavek. On calculation of chamfer distance and lipschitz covers in digital images. In International Conference on Stereology, Spatial Statistics and Stochastic Geometry, Prague (CZ), 2006, June 2006.

10. W.Q. Yan and M.S. Kankanhalli. Colorizing infrared home videos. In 2003 In- ternational Conference on Multimedia and Expo, 2003. ICME '03. Proceedings., volume 1, pages I-97-100 vol.1, July 2003.

11. L. Yatziv and G. Sapiro. Fast image and video colorization using chrominance blending. IEEE Transactions on Image Processing, 15(5):1120-1129, May 2006.

12. Z. Zhang, H. Cui, H. Lu, R. Chen, and Y. Yan. A colorization method based on fuzzy clustering and distance transformation. In 2nd International Congress on Image and Signal Processing, 2009. CISP '09, pages 1-5, Oct 2009.

13. Y. Zhao, L. Wang, W. Jin, and S. Shi. Colorizing biomedical images based on color transfer. In IEEE/ICME International Conference on Complex Medical Engineering, 2007. CME 2007, pages 820-823, May 2007.

14. Z. Zhen, G. Yan, and M. Lizhuang. An automatic image and video colorization algorithm based on pattern continuity. In 2012 International Conference on Audio, Language and Image Processing (ICALIP), pages 531-536, July 2012.

Content Based Image Retrieval in the Context of Alzheimer's Disease

Katarina Trojacanec[1], Ivan Kitanovski[1], Ivica Dimitrovski[1], Suzana Loshkovska[1], for the Alzheimer's Disease Neuroimaging Initiative[*]

[1] Ss. Cyril and Methodius University
Faculty of Computer Science and Engineering, Skopje
Rugjer Boshkovik 16, PO Box 393,
Skopje, Macedonia
{katarina.trojacanec, ivan.kitanovski, ivica.dimitrovski, suzana.loshkovska}@finki.ukim.mk

Abstract. The paper makes an overview of the current status and challenges in 3D Content Based Medical Image Retrieval (CBMIR) with respect to the Alzheimer's Disease (AD). The basic concepts including volume of interest (VOI) selection/detection/localization, feature extraction and similarity measurements are discussed and analysed regarding AD. The possibilities for improvements and challenges of the CBMIR applied to AD are highlighted.

Moreover, a model for CBIR of MRI aimed to be used in the context of AD is proposed. It contains the basic modules that enable the CBIR mechanism such as segmentation, feature extraction and similarity measurement modules. It has additional modules aimed to address the progression of the disease which is crucial for AD. The ultimate goal is to answer the question of retrieving the volumes/cases with the same/similar progress of the condition/disorder/disease with the progress for a given patient regarding AD which is a big challenge not yet exhaustively solved in this area. The model is applied to MRIs provided by the Alzheimer's Disease Neuroimaging Initiative (ADNI). The feature extraction is based on the quantitative measurements obtained from the VOIs assumed to be relevant for AD. To evaluate the retrieval performance, mean average precision (MAP) was used. According to the results, the best MAP of 0.71% is reached when volumes of the subcortical structures are used as feature representation.

Keywords: 3D-CBMIR, Alzheimer's Disease, ADNI, VOI, periodical scans, segmentation, feature extraction, similarity measurements.

1   Introduction

According to the rapid development of the medical imaging techniques, huge amount of 3D or even more dimensional medical images are acquired and stored nowadays in the medical and research centers. They provide plentiful clinical information, coherent and collective view, thus becoming one of the most powerful diagnostic techniques [1]. This leads to a vast amount of 3D medical imaging collections available nowadays that need to be efficiently and precisely searched. The systems that provide efficient retrieval, referred to as Content-Based Image Retrieval (CBIR) systems might have significant importance for clinical, research, and/or educational purposes even though their searching capabilities for medical purposes are one of the biggest challenges among researchers in this domain [1], [2], [3], [4], [5].

Although the researchers in this domain have published several improvements, the search capabilities of medical CBIR still remain questionable in respect to their effectiveness and efficiency [6]. The reason is mostly related to the specific nature of the medical images and addressing it in the CBIR context. For example, Magnetic Resonance Imaging (MRI) is characterized by limited resolution, intensity inhomogeneity, noise, partial volume effects, leading to geometrical inaccuracies that can directly affect the CBIR performance. This is the reason that leads the research of this domain in the direction of specialized CBIR with respect to the application domain rather than general purpose CBIR [1], [2], [3], [4], [5], [7]. Research has been performed towards improvement of 3D Content Based Medical Image Retrieval (3D CBMIR) in two possible research directions/goals [2]: clinical decision support [8] and educational purposes [9].

The application domain considered in this research is Alzheimer's Disease (AD). It is closely related to the structural changes in the brain [10]. 3D MRI with high spatial resolution is considered as proper technique for visualization of subtle anatomical changes that is useful for detection of brain atrophy due to the disease. This is very useful for detecting the pre-disease conditions and monitoring the progression. One important aspect from the clinical point of view is the progression of the patients with mild cognitive impairment (MCI) which is very often a transitional stage between the cognitive decline related to the normal ageing and that related to AD [11].

In this paper, we make an overview of the basic aspects of 3D CBMIR from the point of view of AD application domain. In that context, we review the current status and widely used methods regarding selection/detection/localization of the volume of interest (VOI), feature extraction, and similarity measurements and discuss their suitability, highlight the challenges and future directions to improve the CBMIR with respect to AD. Currently, such a comprehensive overview in this domain does not exist. Moreover, a model for content based retrieval of MRI for addressing AD-related issues is proposed and applied to MRI provided by ADNI database.

The paper is organized as follows. Section 2 provides an overview of the segmentation in the context of AD. The current status of the feature extraction algorithms for images related to AD is summarized in Section 3, while the similarity measurements are discussed in Section 4. Section 5 describes the proposed model for 3D-CBMIR in the context of AD and highlights the challenges and future directions, while Section 6

provides details about the application of the model to ADNI database. Concluding remarks are given in the Section 7.

2 Segmentation

The application of the segmentation to the medical domain is very critical, due to the complexity and variability of the region/volume of interest. Some of the reasons are normal anatomic variation, post-surgical anatomic variation, vague and incomplete boundaries, inadequate contrast, artifacts and noise [12]. Hence, an appropriate and distinctive segmentation method is required [13]. Taking into account that there is no single segmentation technique that can produce satisfactory results for all application domains [14], the segmentation methods are optimized into different directions considering the medical application domain: (1) specific imaging modalities, such as MRI, CT, etc.; (2) specific anatomical structures, such as the brain, the lungs etc.; (3) specific disease/abnormality, such as brain tumor, brain atrophy, etc.

The segmentation is very important for localized monitoring of the changing state of a specific body part that usually reflects the progression of the disease/abnormality, from one side, or the normal state change, from the other side. This is very beneficial from the point of view of AD. Wide research has been performed to detect and highlight possible indicators for AD, and subsequently to analyze their statistical dependence with respect to the disease. Most of them refer to the brain anatomical structures and include cortical thickness [15], [16], [17], ventricular structures [15], [18], hippocampus [17], [19], [20], [21], amygdala [21]. The delineation of such anatomical structures is a key component of further image analysis and interpretation [22].

2.1 State-of-the-Art in AD application domain

Regarding AD, there is plenty of research based on making analysis of VOIs, relevant for detecting anatomical changes related to or imposed by AD. They usually include hippocampus, amygdala, ventricular structures, and brain cortex. Several studies focus on a segmentation of these structures. Multi-atlas segmentation framework used for segmentation of thalamus, caudate, putamen, pallidum, hippocampus and amygdala is proposed in [23]. An improved version of this method used for hippocampus segmentation the authors describe in [21]. Other methods for hippocampus extraction are proposed in [24], [25], [26], [27]. The anatomical segmentation of structural images of the human brain, leading to 83 regions is provided in [28].

Meaningful number of studies is using software tools and packages for the segmentation of the brain structures which are valuable indicators for AD [15], [18], [20], [29], [30]. Among them are: FreeSurfer software package [31] used for cortical and subcortical segmentation [15], [16], Brain Ventricular Quantification (BVQ) software [32] for ventricular segmentation [15], [17], Statistical Parametric Mapping (SPM) software package for White Matter (WM), Grey Matter (GM), and Cerebrospinal Fluid (CSF) segmentation [29], [30], Automatic Lateral Ventricle delIneatioN (ALVIN) for lateral ventricle delineation [33], and FIRST provided by FSL [34].

Considering the segmentation techniques widely used in the literature in this domain, semi- and fully-automated algorithms are required. Some remarkable examples of semi-automated methods that require some user input/interaction include deformable template approach for elastic deformation of the hippocampal model [35], and active contour method augmented by a priori shape information [36]. The application of fully automated methods is also widely spread [11], [19], [24], [25], [28]. They are faced with the lack of "gold standard" for evaluating the segmentation results [37].

3 Feature Extraction

From the point of view of AD application domain, two directions might be considered regarding feature extraction. The first one is the standard procedure, usually used in the traditional CBIR through which image features are derived from the visual cues contained in the image [6] extracted from the Volume of Interest (VOI). Extraction of robust and representative visual features in this application domain is critical and challenging task. Although the research has shown that the more visual features are used the better the results are, in the context of 3D or more dimensional images this might be a problem in same time, due to the significantly increased dimensionality of the feature representation. This imposes the necessity for optimization [7].

It should be noticed that in general, the relative importance of features will vary across modality and disease targets. [6]. Thus, the selection of the feature extraction algorithms should be target oriented rather than generic. Regarding AD, several investigations have been performed to evaluate feature extraction methods. For instance, intensity histograms, local binary pattern and gradient magnitude histograms are used to generate feature vector for the middle slice for subsequent usage in automated diagnosis of AD [38]. Discrete Cosine Transform (DCT), Daubechie's Wavelet Transform (DWT) and Local Binary Patterns (LBP) are applied on 2D bases on a selected by radiologists subset of slices [39]. Laguerre Circular Harmonic Functions expansions enabling capturing the local image patch structure are also used for feature extraction [40]. The Bag-of-Visual-Words approach is then followed on a specific region (hippocampus). Slice by slice analysis is performed in this research too.

All of these studies are performed only on one slice or slice by slice manner, which means excluding possibly significant information that might be extracted from volumetric data. We suggest that 3D extensions of these descriptors might be more suitable and more comprehensive. For instance, a suited version of grey level co-occurrence matrices to 3D context, 3D Local Binary Patterns (3D LBP), 3D Wavelet Transforms (3D WT), and 3D Gabor Transforms (3D GT) [1]. Another descriptor that may also be powerful in the context of AD is 3D SIFT descriptor introduced in [41].

On the other side, wide range of studies have been performed to detect and highlight possible indicators for AD, and to analyze their statistical dependence with respect to the disease [15], [18], [19], [20], [21]. This leads to another possible direction for feature extraction, which means using measurements for the selected brain structures as features. Some of them are used by the researchers for distinguishing or automatically label/classify patients as AD, MCI, or healthy [11], [42], [43]. While the

first direction means direct visual information description, the second one uses the visual information only to delineate the valuable (for AD) brain structures on the bases of which the quantitative measurements are obtained.

4   Similarity Measurements

Based on the empirical estimates of the feature distribution, different similarity measurements have been used to make a similarity comparison: Minkowski-Form distance (a generic form of the Euclidean distance), Mahalanobis distance, quadratic form distance, proportional transportation distance, earth mover's distance, Kullback-Leibler divergence and Jeffrey divergence, etc. [44]. If the graph-based representation of the features is used rather than vector-based, special computational methods for similarity assessment, referred to as graph matching, are required. There are cases, especially for diagnostic purposes, where subtle geometrical differences between structures are very important and the similarities are defined through the notion of elastic deformations required to transform one shape into another are useful. Another, very promising approach includes statistical classifiers that classify new instances using high-level information extracted from the training set of instances. This approach is very problem oriented. Choosing a metric to measure the similarity has direct influence on the CBIR performance. It is dependent on the descriptor type [6].

5   Model for CBMIR in the Context of Alzheimer's Disease

In this section, a model for CBIR aimed to be used in the context of Alzheimer's Disease is described. It is depicted on Fig. 1. While containing modules that come from the general CBIR architecture, adapted to AD, the model also contains modules that are incorporated to treat the progression of the disease problem. Each module could be separately improved to solve its specific problem, giving a proper impact to the whole model, and reducing error propagation through the connected modules.

Preprocessing Module. The first module is related to the preprocessing aimed to provide image corrections if needed. Several steps/techniques are notable here: partial volume effect correction [26], gradient non-linearity correction, correction for image intensity inhomogeneity, and bias field correction [45].

Segmentation Module. Regarding the segmentation module, techniques for segmentation of VOIs adapted to the MRI image modality and AD application domain should be used. Several challenges are identified in this context: (1) robustness to the difficulties imposed by image acquisition process; (2) capability to properly delineate the relevant structure; (3) decreased user/expert influence by using semi- or fully automated algorithms; (4) degree of the domain knowledge usage; (5) complexity and processing time.
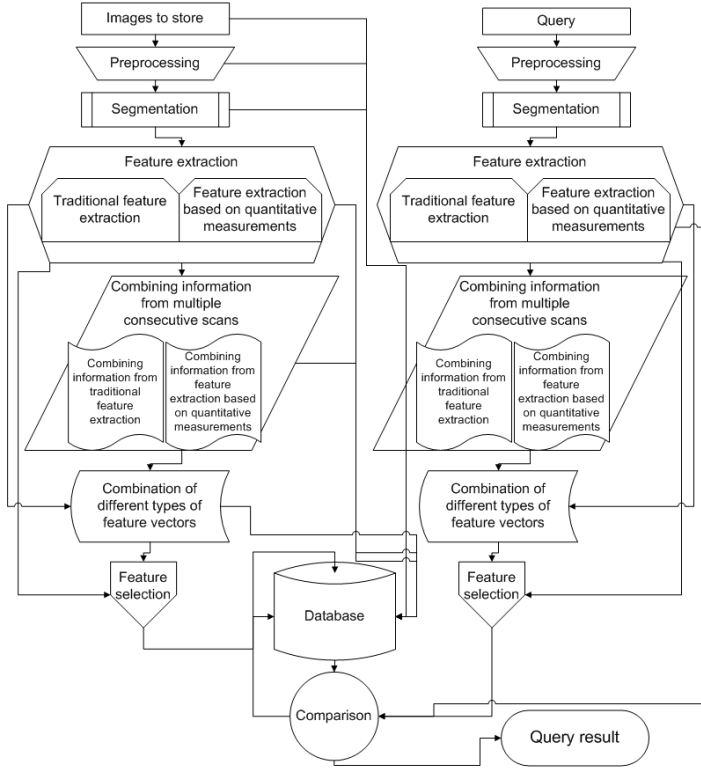
Fig. 1. Model for CBIR intended to be used in the context of Alzheimer's Disease

Feature Extraction Module. This module is comprised of two sub-modules. The first uses the feature extraction procedure which is characteristic for the traditional CBIR, while the second is based on the quantitative measurements derived from the valuable markers for AD.

The first sub-module is used to extract information from the visual cues contained in an image. The techniques in this sub-module should be focused on precise information extraction directly from the volume content. Taking into consideration the spatial relationships in this context is an important challenge that should be addressed.

The second sub-module regarding the feature extraction is based on the quantitative measurements obtained from the VOIs that are assumed to be valuable indicators for AD. Taking the results of the segmentation, the calculation/estimation of the significant measurements is performed in this sub-module, generating a feature vector comprised of these quantitative measurements.

Conceptually, the difference between the two sub-modules is that using the features in the retrieval process provided by the former would lead to answering questions of type "find all images that have similar visual properties to the query image/VOI", while using the features generated by the later, would mean answering question of type "find all images that have similar structural changes to the query".

Module for Combining Information from Multiple Consecutive Scans. This module is aimed to address the case when multiple scans, obtained from different time points, are available for a given patient. Its goal is to extract/combine information from the available scans of the patient. This module is very important because it enables incorporation and addressing the change/progression of the patient's condition/possible disease. Two approaches are possible here:

- To use the information provided by the first sub-module of the feature extraction module for all available scans and make a combination. This can be accomplished by using a simple concatenation or an algorithm that will create a feature vector prototype from the feature vectors generated for each scan for a given patient.
- To use the information provided by the second sub-module of the feature extraction module for all available scans and make a combination. Similarly as in the previews case, the combination might be performed by simple concatenation, or an algorithm that will reflect the change of each feature/indicator over the time or eventually produce a prototype from the quantitative measurements obtained at each time point for a patient.

Additionally, this module should be able to handle the missing data. It is highly possible that different series of scans are available for different cases/patients. This will lead to different dimension of the feature vectors if any kind of concatenation is used or missing information if any other algorithm for combination is used. One simple solution that might be used in this case is to consider only those subjects who have all scans for the examined time points and exclude the (possibly large group of) patients with no sufficient information available. Another solution which arises to overcome the drawback of excluding some cases is creating an artificial sample for the missing scan. This might be accomplished in two different ways. The first is making an artificial model of the brain/VOIs which as a representation of the missing scan on the bases of some kind of interpolation regarding the analysis of anatomical variation over the time. The second way is deriving the information of the missing scan directly from the feature vectors. This can be conducted using the information obtained from the patients who have full series of scans to predict the missing information.

Module for Combination of Different Types of Feature Vectors. This module is optional and provides a combination of the different types of feature vectors, those derived in a traditional way (obtained in the first sub-module of the feature extraction module) or those comprised of the quantitative measurements derived from the valuable indicators (obtained in the second sub-module). This might be helpful for addressing patients with pre-disease condition such as MCI. In this case, volume change of the hippocampus for example is not as significant marker as in AD. To overcome this, including information about hippocampal shape is suggested [19].

Feature Selection Module. It is very common in this domain to obtain feature vectors that are high dimensional. The feature selection module tends to reduce the dimensionality of the feature vector by analyzing the importance/impact of the features.

Similarity Measurements Module. Next is the similarity measurements module which provides comparison between the query and all volumes stored in the database. It is highly dependable of the feature extraction process. It is very important to choose appropriate similarity/dissimilarity measurement and comparison between the query unit and all units in the database.

Discussion. Taking into consideration the aforementioned modules, the model should be able to find the cases with the same/similar progress of the condition/disorder/disease with the progress of a given patient. For example, a few queries we found challenging in the context of AD:

- Find all cases with progression from MCI to AD similar to the query one
- Find all cases with MCI that have (not) progressed to AD for the examined period similar to the query one
- Find all cases with the similar progression of AD to the query one

The model is designed to be able to answer the questions that might be helpful from different aspects, including diagnosis and prognosis support, educational purposes, completion of the general image about the patient's state in the case of missing one or more scans of his/her study, proposing the treatment and/or monitoring the patient's reaction on particular treatment.

6   Experimental Setup

6.1   Dataset

The images for this research were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies, and non-profit organizations as a $60 million, 5-year public–private partnership. Conducting examination on whether the combination of serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, such as cerebrospinal fluid (CSF) markers, APOE status and full-genome genotyping via blood sample, as well as clinical and neuropsychological assessments can be used to measure the progression of mild cognitive impairment (MCI) and Alzheimer's Disease (AD) has been the primary goal of ADNI. Determination of sensitive and specific markers of very early AD progression is aimed to assist the development of new treatments, to enhance the process of monitoring treatments effectiveness, and to reduce the time and cost of clinical trials. The dataset includes cognitively normal individuals, adults with early or late MCI, and people with early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2, and ADNI-GO. For up-to-date information, see http://www.adni-info.org.

6.2 Implementation details

Going down through the pipeline described in Section 5, this subsection includes implementation details of the modules used for the purpose of this research.

With the aim to reduce image nonidealities, the ADNI data are available with different levels of preprocessing, including gradient non-linearity and intensity non-uniformity correction, phantom-based distortion correction, and histogram peak sharpening algorithm for bias field correction [45]. Taking this into consideration, no additional preprocessing steps are conducted in the preprocessing module.

Considering the segmentation module, the FreeSurfer software package version 5.1.0 was used, enabling cortical and subcortical segmentation. The subcortical structures of interest for this research include left and right lateral ventricle, third and fourth ventricle, hippocampus, amygdala, as well as the cortical structures.

From the point of view of the feature extraction, the second sub-module of the feature extraction module takes place. The representation of the information extracted from the MRI volumetric data is based on the quantitative measurements of the structures assumed to be mostly affected by AD obtained by using FreeSurfer. The feature vector is constructed from the left and the right lateral ventricle volume, the third and the fourth ventricle volume, the volume of the left and right hippocampus, left and right amygdala, as well as the cortical thickness of the separate cortical structures from the left and right hemisphere (34 cortical regions for each hemisphere). This research is conducted on several cases where the feature vector consists of:

- Volumes of the subcortical structures (left and right lateral ventricle, third and the fourth ventricle, left and right hippocampus, left and right amygdala) with total length of eight features
- Cortical thickness of the cortical regions of the left and right hemisphere (34 regions for each hemisphere), with total length of 68 features
- Both, volumes of the subcortical structures, and cortical thickness, leading to 76 features in total

It should be noted that the dimensionality in all three cases is relatively small in comparison to the traditional descriptors, e.x. 13312 features for 3D Gray Level Cooccurrence Matrices, 1920 for 3D Wavelet Transforms, 9216 for Gabor Transforms, and 11328 for 3D Local Binary Patterns for one volume [3].

Due to the differences in the head size between individuals, an important question arises: what normalization approach to be used. It is usually suggested that the volumes of the subcortical structures should be normalized by the Intracranial Volume (ICV). Considering the cortical thickness, this question does not have clear answer, although research has been conducted normalizing it by the mean cortical thickness and it is concluded that it does not have significant influence [46], [47]. We provide results with and without normalization for each form of the feature vector, to illustrate its influence on the retrieval performance, i.e. with and without normalization by the ICV for the volumes of the subcortical structures, as well as with and without normalization by the mean cortical thickness of the left and right hemisphere respectively.

It should be noticed that he modules for combining information from multiple consecutive scans and for combining different types of feature vectors, as well as the feature selection module are not used in this research and are planned to be further researched in our future work.

As a similarity measurement, Euclidean distance is used as it is considered as appropriate metric to the feature representation.

6.3   Results

For this research, the baseline images obtained using 3T scanners from ADNI database were used. Two categories of patients were included for the purpose of the research: patients with AD (33 subjects), and normal controls, NL (47 subjects). Because of the small number of images in this subset, leave-one-out strategy was performed. This means that each image was used as a query against all other images in the database. To evaluate the retrieval performance, mean average precision (MAP) was used. The retrieved image is considered as relevant if it belongs to the same class as the query (AD or NL). Table 1 summarizes the results (on the bases of MAP) for each form of the feature vector with and without normalization.

Table 1. Evaluation of the retrieval performance on the bases of MAP

Feature vector form	MAP
Subcortical volumes – raw (8 features)	0.71
Subcortical volumes – normalized by ICV (8 features)	0.71
Cortical thickness – raw (68 features)	0.62
Cortical thickness – normalized by the mean thickness (68 features)	0.62
Subcortical volumes – raw + Cortical thickness – raw (76 features)	0.66
Subcortical volumes – normalized by ICV + Cortical thickness – raw (76 features)	0.66
Subcortical volumes – raw + Cortical thickness – normalized by the mean thickness (76 features)	0.66
Subcortical volumes – normalized by ICV + Cortical thickness – normalized by the mean thickness (76 features)	0.66

According to the results given in Table 1, we can conclude that using the representation where the feature vector consists of the volumes of the subcortical structures leads to the best retrieval results regarding the MAP. The combination of the subcortical volumes and the cortical thickness leads to smaller value of MAP, while the smallest MAP is obtained when the feature vector is constructed only by the cortical thickness values of the cortical regions. The explanation for such negative influence of the cortical thickness on the retrieval performance is that the cortical structures are only a few voxels thick [48]. This can result to errors of the measuring algorithms which can be subsequently propagated to the retrieval process.

Additionally, it can be concluded that for the subset used for the research in this paper, the normalization factor (ICV for the subcortical volumes and mean thickness for the cortical thickness of the cortical regions) does not have any influence.

To be able to compare the results to the researh performed on the same subset of ADNI dataset in [40], we also provide the curves of average precision at the first N retrieved scans (Fig. 2).
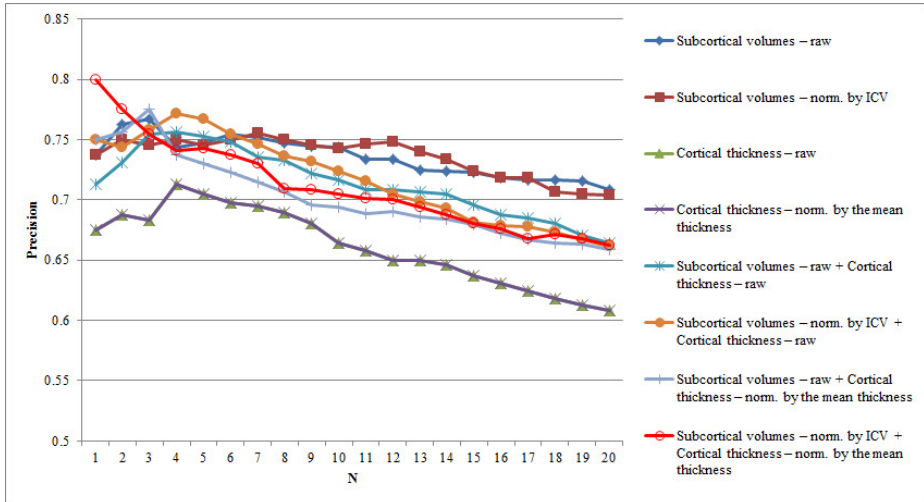
Fig. 2. Precision at N for the subset of ADNI dataset containing AD and NL subjects (obtained using 3T scanners)

According to the results reported in [40] for this subset, the average precision reaches 0.74 (at N=1). In our case, we obtain the same average precision at the same point in the case of feature vector comprised of the volumes of the subcortical structures (eight features). However, the best average precision at this point is obtained when the feature vector consists of the volumes of the subcortical structures normalized by ICV and the cortical thickness normalized by the mean cortical thickness (76 features). This means that the method for using measures of the brain structures used in this paper reaches the same average precision at N=1 with [40] using the feature vector of only eight features, and reaches even better average precision of 0.8 at the same point in the case of the concatenated feature vector of 76 features. According to this, the results obtained using selected stages of the proposed model in this research give promising results.

7 Conclusion

A comprehensive overview of the current status of the 3D CBMIR with respect to Alzheimer's Disease was provided in the paper. Discussion regarding its crucial concepts such as segmentation of the VOI, feature extraction, as well as the similarity/dissimilarity measurements from the perspective of their suitability to this application domain was given. The most remarkable challenges and the room for improvements were detected and highlighted.

Additionally, a model for 3D CBIR of MRI aimed to be used in the context of AD was proposed. Its ultimate goal is to enable efficient and precise answer of the questions of finding VOIs/images/cases with similar condition to the query one and of retrieving the VOIs/images/cases with the same/similar progress of the condition/disorder/disease with the progress for a given patient regarding AD which is a big challenge not yet exhaustively solved in this area. Answering this questins is valuable for educational purposes, providing completion in the case of missing one or more scans of his/her study, proposing the treatment and/or monitoring the patient's reaction on particular treatment.

Additionally, the model was evaluated on a sub-set of ADNI database. The obtained preliminary results reach MAP of 0.71. We compared the results to those results obtained by other researchers on the same sub-set, where the average precision reaches 0.74 (at the first retrieved image). In this research, the same average precision is obtained at the same point, but with the feature vector comprised of only eight features, and even better average precision of 0.8 in the case of concatenated feature vector of 76 features (at the same level, the first retrieved image).

For future work, we are planning to extent this research including the other modules, and what is most important, treating the progression of the disease. Additionally, evaluation of the model on Q bigger dataset is planned to be performed.

References

1. Xiaohong, G., Qian, Y., Loomes, M., Comley, R., Barn, B., Chapman, A., Rix, J., Hui, R., Tian, Z.: Retrieval of 3D medical images via their texture features. International Journal on Advances in Software , 4(3), 499--509 (2012)
2. Foncubierta, R.A, Müller, H., Depeursinge, A.: Region-based volumetric medical image retrieval. In Proceedings of SPIE 8674, Medical Imaging 2013: Advanced PACS-based Imaging Informatics and Therapeutic Applications, pp. 867406--867406 (2013)
3. Qian, Y., Gao, X., Loomes, M., Comley, R., Barn, B.: Content-based Retrieval of 3D Medical Images. In: 3rd International Conference on eHealth, Telemedicine, and Social Medicine (eTELEMED 2011), pp. 7--12 (2011)
4. Depeursinge, A., Zrimec, T., Busayarat, S., Müller, H.: 3D lung image retrieval using localized features. In SPIE Medical Imaging, International Society for Optics and Photonics, pp. 79632E--79632E (2011)
5. Agarwal, M., Mostafa, J.: Image Retrieval for Alzheimer's Disease Detection. In Medical Content-Based Retrieval for Clinical Decision Support, pp. 49--60, Springer, Berlin Heidelberg (2010)
6. Akgül, C.B., Rubin, D.L., Napel, S., Beaulieu, C.F., Greenspan, H., Acar, B.: Content-based image retrieval in radiology: current status and future directions. Journal of Digital Imaging, 24(2), 208--222 (2011)
7. Simonyan, K., Modat, M., Ourselin, S., Cash, D., Criminisi, A., Zisserman, A.: Immediate ROI search for 3-D medical images. In: Medical Content-Based Retrieval for Clinical Decision Support, pp. 56--67, Springer Berlin Heidelberg (2013)
8. Osechinskiy, S., Kruggel, F.: PDE-based reconstruction of the cerebral cortex from MR images. In: IEEE International Conference on Engineering in Medicine and Biology (EMBC'10), pp. 4278--4283 (2010).
9. Rosset, A., Muller, H., Martins, M., Dfouni, N., Vallée, J.-P., Ratib, O.: Casimage project - a digital teaching files authoring environment, Journal of Thoracic Imaging 19(2), 1--6 (2004)
10. Agarwal, M., Mostafa, J.: Content-based image retrieval for Alzheimer's disease detection. In Content-Based Multimedia Indexing (CBMI), 2011 9th International Workshop on, pp. 13--18 (2011)
11. Gray, K.R., Aljabar, P., Heckemann, R.A., Hammers, A., Rueckert, D.: Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. NeuroImage, 65, 167--175 (2013)
12. Shijun, W., Summers, R. M.: Machine learning and radiology. Medical image analysis 16(5), 933--951 (2012).
13. Zhen, M., Manuel, J. Tavares, R.S, Jorge, R.N. Mascarenhas, T.: A review of algorithms for medical image segmentation and their applications to the female pelvic cavity. Computer Methods in Biomechanics and Biomedical Engineering 13(2), 235--246 (2010)
14. Farag, A.A., Ahmed M.N., El-Baz A, Hassan, H.: Advanced segmentation techniques. In Handbook of biomedical image analysis, pp. 479-533, Springer, US (2005)
15. Moore, D.W., Kovanlikaya, I., Heier, L.A., Raj, A., Huang, C., Chu, K.W., Relkin, N.R.: A pilot study of quantitative MRI measurements of ventricular volume and cortical atrophy for the differential diagnosis of normal pressure hydrocephalus. Neurology research international 2012 (2011)
16. Yuan, L., Wang, Y., Thompson, P.M., Narayan, V.A., Ye, J.: Multi-source feature learning for joint analysis of incomplete multiple heterogeneous neuroimaging data. NeuroImage, 61(3), 622--632 (2012)

17. Sabuncu, M.R., Desikan, R.S., Sepulcre, J., Yeo, B.T.T, Liu, H., Schmansky, N.J., Reuter, M. et al.: The dynamics of cortical and hippocampal atrophy in Alzheimer disease. Archives of neurology 68(8), 1040--1048 (2011)

18. Nestor, S.M., Raul, R., Michael, B., Matthew, S., Vittorio, A., Jennie, L.W., Jennifer, F., Robert, B.: Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database. Brain 131(9), 2443--2454 (2008)

19. Gerardin, E, Gaël, C., Marie, C., Rémi, C., Béatrice, D., Ho-Sung, K., Marc, N. et al.: Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. Neuroimage 47(4), 1476--1486 (2009)

20. Leonardo, I.: Atrophy Measurement Biomarkers using Structural MRI for Alzheimer's Disease. The 15th Int. Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (2012)

21. Lötjönen, J., Robin W., Juha K., Valtteri J., Lennart T., Roger L., Gunhild W., Hilkka S., Daniel R.: Fast and robust extraction of hippocampus from MR images for diagnostics of Alzheimer's disease. Neuroimage 56(1), 185--196 (2011)

22. Leemput, V.K. Vandermeulen, D., Maes, F., Srivastava S, D'Agostino E., Suetens P.: Model-Based Brain Tissue Classification. In: Handbook of Biomedical Image Analysis, pp. 1--55. Springer, US (2005)

23. Lötjönen, J.M., Wolz, R., Koikkalainen, J.R., Thurfjell, L., Waldemar, G., Soininen, H., Rueckert, D.: Fast and robust multi-atlas segmentation of brain magnetic resonance images. Neuroimage, 49(3), 2352--2365 (2010)

24. Chupin, M., Gérardin, E., Cuingnet, R., Boutet, C, Lemieux, L., Lehéricy, S., Benali, H., Garnero, L., Colliot, O.: Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. Hippocampus 19(6), 579—587 (2009)

25. Morra, J. H., et al.: Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer's disease mild cognitive impairment, and elderly controls. Neuroimage, 43(1), 59--68 (2008)

26. Chupin, M., Mukuna-Bantumbakulu, A.R., Hasboun, D., Bardinet, E., Baillet, S., Kinkingnehun, S., et al.: Anatomically constrained region deformation for the automated segmentation of the hippocampus and the amygdala: method and validation on controls and patients with Alzheimer's disease. NeuroImage. 34, 996--1019 (2007)

27. Chupin, A., Hammer, A., Liu, R.S., Colliot, O., Burdett, J., Bardinet, E., Duncan, J.S., Garnero, L., Lemieux, L.: Automatic segmentation of the hippocampus and the amygdala driven by hybrid constraints: method and validation. Neuroimage. 46(3), 749--761 (2009)

28. Heckemann, R.A., Keihaninejad, S, Aljabar, P., Gray, K.R. , Nielsen, C, Rueckert, D., Hajnal, J.V., Hammers, A.: Automatic morphometry in Alzheimer's disease and mild cognitive impairment." Neuroimage 56(4), 2024--2037 (2011)

29. Cataldo, R., Agrusti, A., De Nunzio, G., Carlà, A., De Mitri, I., Favetta, M., Quarta, M., Monno., L., Rei., L., Fiorina, E.; Alzheimer's Disease Neuroimaging Initiative: Generating a minimal set of templates for the hippocampal region in MR neuroimages. Journal of Neuroimaging 23(3), 473--483 (2013)

30. Casanova, R., Hsu, F.C., Espeland, M.A., Alzheimer's Disease Neuroimaging Initiative: Classification of structural MRI images in Alzheimer's disease from the perspective of ill-posed problems. PLoS One. 7(110), e44877 (2012)

31. FreeSurfer, http://freesurfer.net/

32. Accomazzi, V., Lazarowich, R., Barlow, C.J., Davey, B.. inventors; Cedara Software Corp., assignee.: Image region segmentation system and method. US Patent Application 20050031202 (2005)

33. ALVIN lateral ventricle segmentation, https://sites.google.com/site/mrilateralventricle/

34. FSL, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/

35. Shen, D., Moffat, S., Resnick, S., Davatzikos, C.: Measuring size and shape of the hippo-campus in MR images using a deformable shape model. Neuroimage. 15(2), 422--434 (2002)

36. Ferrarese, P. F., Menegaz, G.: Performance Evaluation in Medical Image Segmentation. Current Medical Imaging Reviews 9(1), 7--17 (2013)

37. Bauer, S., Wiest, R., Nolte, L. P., Reyes, M.: A survey of MRI-based medical image anal-ysis for brain tumor studies. Physics in medicine and biology. 58(13), R97 (2013)

38. Akgül, C. B., Ünay, D., Ekin, A.: Automated diagnosis of Alzheimer's disease using image similarity and user feedback. In Proceedings of the ACM International Conference on Im-age and Video Retrieval, pp: 34 (2009)

39. Agarwal, M., Mostafa, J.: Content-based image retrieval for Alzheimer's disease detection. In Content-Based Multimedia Indexing (CBMI), 2011 9th International Workshop on, pp: 13--18 (2011)

40. Mizotin, M., Benois-Pineau, J., Allard, M., Catheline, G.: Feature-based brain MRI re-trieval for Alzheimer disease diagnosis. In Image Processing (ICIP), 19th IEEE Interna-tional Conference on, pp: 1241--1244 (2012)

41. Scovanner, P., Ali, S., Shah, M.: A 3-dimensional sift descriptor and its application to ac-tion recognition. In: Proceedings of the 15th International Conference on Multimedia ACM, pp. 357--360 (2007)

42. Gray, K. R., Wolz, R., Heckemann, R. A., Aljabar, P., Hammers, A., Rueckert, D.: Multi-region analysis of longitudinal FDG-PET for the classification of Alzheimer's disease. NeuroImage, 60(1), 221--229 (2012)

43. Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M. O., Chupin, M., Benali, H., Colliot, O.: Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. Neuroimage 56(2), 766--781 (2011)

44. Gao, X., Y. Qian, R. Hui, M. Loomes, R. Comley, B. Barn, A. Chapman, J. Rix.: Texture-based 3D image retrieval for medical applications. IADIS e-Health 2010, pp. 29--31 (2009)

45. Wyman, B. T., Harvey, D. J., Crawford, K, Bernstein, M. A., Carmichael, O., Cole, P. E., Crane, P. K. et al.: Standardization of analysis sets for reporting results from ADNI MRI data. Alzheimer's & Dementia 9(3) 332--337 (2013)

46. Velayudhan, L., Proitsi, P., Westman, E., Muehlboeck, J. S., Mecocci, P., Vellas, B.,et al.: Entorhinal cortex thickness predicts cognitive decline in Alzheimer's disease. Journal of Alzheimer's Disease, 33(3) 755-766 (2013)

47. Westman, E., Aguilar, C., Muehlboeck, J.S., Simmons, A.: Regional magnetic resonance imaging measures for multivariate analysis in Alzheimer's disease and mild cognitive im-pairment. Brain topography, 26(1), 9--23 (2013)

48. Li, Y., Wang, Y., Wu, G., Shi, F., Zhou, L., Lin, W., Shen, D:. Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features. Neurobiology of aging, 33(2) 427--e15 (2012)

Models and techniques to evaluate system and software reliability based on software reliability engineering methodology
Case study: Attitude Control System

Orges Çiço[1], Zamir Dika[2]

[1]Economic Faculty, University of Tirana,
Rr. Elbasanit, Tirana, Albania
[2]Faculty of Contemporary Sciences and Technologies
South East European University, Tetovo, FYR of Macedonia
orges.cico@gmail.com, z.dika@seeu.edu.mk

Abstract. The main goal of this paper is the evaluation of the state of the art techniques developed to estimate software reliability and prove their effectiveness on a real case study. The primary issue with Software Reliability Engineering (SRE) derives from its feasibility on real life projects, in terms of software cost, complexity, application in various environments and involved staff qualification. Although, hardware modeling has been well understood during the past decades, software still offers some challenges due to its not material nature. Many models have been proposed and different techniques adopted from hardware fault tolerant techniques, but they have yet to be proven effective on real case studies, involving critical software systems. Thus, reliability testing and modeling of the Attitude Control System (ACS) of a nanosatellite represents the problem on which this paper is focused. The methodology followed was based on the well recognized and established Software Reliability Engineering Process.

1   Introduction

Technology has reached its greatest boom during the last decades, especially in computer-based systems. They, together with the software operating on them have become more and more present in our daily lives. The way society is evolving shows how much it relies on this systems in many different aspects, such as home appliances, personal technological devices, transportation vehicles, telecommunications, airplanes, money, teaching, medical devices, power sources and so on. Criticality of the various systems imposes different requirements of software reliability and quality.

Nowadays, technologies and sciences rely heavily on systems running highly

qualitative software in order to bring higher knowledge impact and what is more important, to avoid critical system failures. It is worth mentioning that failure-free software operation is crucial for medical devices, banking and commerce, nuclear power plants, airplanes, aeronautical devices and space satellites, etc. Unfortunately, the size and complexity of the software is closely related to the complexity of the systems for which it has been developed.

The last half century of technological human history has pointed out an exponential growth of the previously mentioned attributes of software systems [1], especially in the fields of telecommunications, business and transportation. This leaves us with the challenge to provide these technologies with safer and more reliable software, since its failure may even imply catastrophic consequences to human lives [2]. Experience has proved that, despite the advanced processes and tools throughout which nowadays software is being developed, analyzed and tested, it is very dangerous to assume it to be error-free. Therefore, emerges the necessity to accurately model [3], [4], [5], [6] and predict its future behavior. We can than equip our implementation with the necessary means and techniques to circumvent errors and thus annoying failures [7], [3], [8].

2   Background

Dependability according to [2] is defined as the trustworthiness of a computer system such that reliance can justifiably be placed on the service it delivers. The service delivered by a system is its behavior as it is perceived by its users. A user is considered to be another system, either physical or human, that interacts with the former throughout the service interface. The delivered service is considered as a correct one when it implements the required system function. The function of a system is what the system is intended to do, and is described by the functional specification. When the service delivered by the system is incorrect the system itself is considered to be out aged. However, measures can be taken to restore the system, so that it can still deliver a correct service.

Fig.1 shows all the elements that influence and compose the dependability of a system. One of the crucial attributes to be considered during our research work is the system reliability; this represents the scope of this paper work.

3   State of the art and related works

It is convenient to determine the errors remaining within a software system, but in order to model and predict its future behavior consistent data need to have been gathered from the past failures. During the last 35 years different models and techniques have been developed in order to measure the current and future reliability of the software systems. Although it is impossible to construct a totally failure free software, it is important to either identify an existing model or propose a new one that best describes the probability of its future failures and thus giving to the final user the adequate quality assurance. Several software reliability models have been proposed

and used by different tools for various case studies. In general, two main types of software reliability models can be identified: Deterministic Reliability Models and Probabilistic Reliability Models.
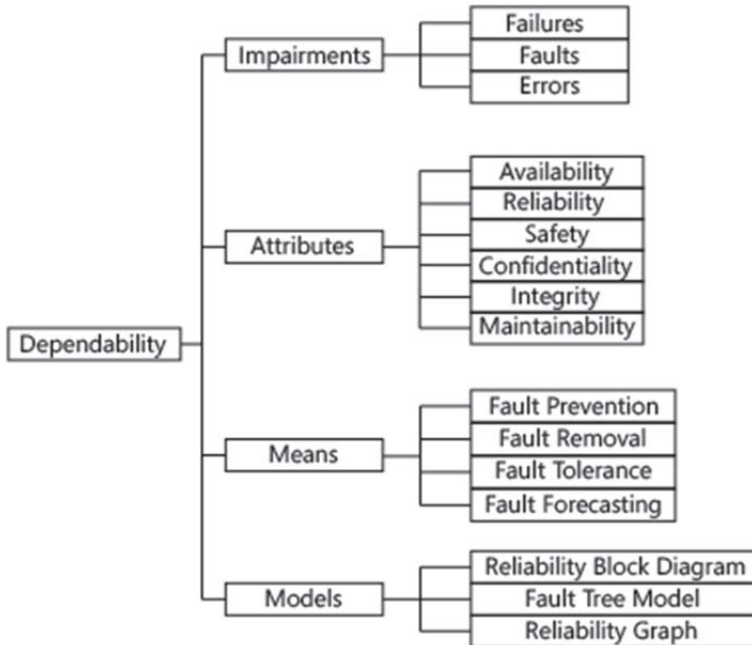
Fig. 1. Dependability: Impairments, Attributes, Means, Models

The first group of models relies on totally deterministic data such as the number of machine instructions, errors, operators and operations within the source code of a program, to provide measures which characterize the software. None of these measures has any probabilistic nature. The most well known deterministic models are: Halstead's software metric and McCabe's cyclomatic complexity metric

Halstead's software metric is used to estimate the number of errors in the program, whereas McCabe's cyclomatic complexity metric (McCabe 1976) is used to determine an upper bound on the model for estimating the number of remaining software defects. While in the second group of models fault removals and failure occurrences are considered as probabilistic events. According to (Pham 2000a) probabilistic reliability models are classified in:

- Error seeding [9]
- Failure rate [10], [11]
- Reliability growth [12], [13]
- Markov structure
- Time-series
- Non homogeneous Poisson process

Each category contains several different models with different probability distribution function. Our focus is put mainly on probabilistic models since we are primary interested on the future behavior of the software and the prediction of its failures, on random environmental conditions. According to Lyu, [14], models are classified into:

- Time domain: Classification based on calendar or execution time.
- Category: Classification based on the number of failures which is either Finite or infinite.
- Type: Classification based on the distribution of the number of Failures experienced by the specified time.
- Class: Classification considered only for the finite category. It is Based on the functional form of the failure intensity over time.
- Family: Classification considered only for the infinite category. It is based on the functional form of the failure intensity in terms of the expected number of failures experienced.

Most of the models explored in the literature review are commonly adapted to the CASRE (Computer Aided Software Reliability Engineering) tool. The aim of the tool is to fully exploit the different models and adopt the right one so that the reliability prediction is the most accurate possible.

4   Testing methodology and steps

In order to perform load testing we need to define the operations that the software system is going to perform. We will test the software accordingly, for the most critical and with highest probability of occurrence operations.

4.1 Operational Profile development

In this section we will explain in detail how to construct the operational profile and document testing, for both performance and load. There are two main motivations in developing an operational profile:

1. Reflecting in practice the software system usage
2. Guiding the testing

The operational profile usually depends on functionality of the software, user types, duration in time etc. Operations, differently from functions, are determined during the design lifecycle of the software system. We can identify operations from different design elements such as Use Case, Class or State Flow Diagrams. The operational profile represents its operations in two different manners: Tabular and Graphical.
In the tabular representation we provide a list of operations combined with their

probabilities of occurrence.

In order to construct a good operational profile we need to determine all the relevant operational modes. An operational mode represents distinct pattern of software system use or conditions deriving from the environment, that require separate testing due to the likely-hood of producing different failures. A particular operation may appear in different operational modes with different probability of occurrence.

The methodology in developing the Operational Profile follows the steps shown in Fig. 2. First of all the different operational modes for which testing will be performed need to be determined. Operational mode examples are:

- Critical states (such as shutdown)
- Operational conditions (such as software system overload, exhaustion of resources)
- User types or experience
- Time
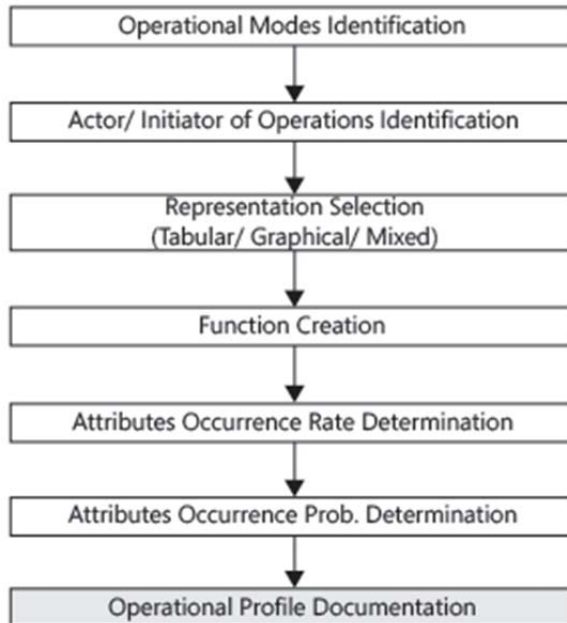
Fig. 2. Operational profile construction steps

Secondly, the possible initiators should be determined. Thirdly, in dependence of the possible attributes for every operation we need to determine if we should use either a tabular or graphical representation of the Operational Profile.

In the following section we will see that the tabular representation is more convenient for our case study. Then an operation list for every operation mode needs to be determined based on several factors such as: implicit or explicit operation, environment variables.

Afterwards, the occurrence rate and occurrence probability of every operation is determined based on field data. The occurrence probability is obtained from:

occurrence probability = occurrence rate of single operation / total occurrence rate

Finally the Operational Profile Documentation is obtained and ready for the testing phase.

4.2 Performing Performance Testing guided from the Operational Profile

Operational testing consists of checking the working of the software in its relevant operational environment. Based on the operational profile, every operation is subdivided into a finite number of runs also known as a group of total occurrence rate run types conceived at development phase. A run is usually considered as the smallest division of work that can be initiated from an external actor. Runs are associated with their input state, usually direct inputs. Runs with identical input states are considered to be part of the same run type. A test case is considered as a partial specification of a run, based on naming the direct inputs and their values. Test cases are independent from the operational mode and can be executed in different operational modes at once, thus discovering different failure behaviors.

There exist mainly three types of testing:

1. Feature Testing - used to check the features provided by the software, throughout single execution of every operation with minimal interaction with other operations.
2. Load Testing - to check performance under maximum workload (e.g. on Databases   Application Servers etc.).
3. Regression Testing - feature test performed after every bug has been fixed.

However, in this paper we will focus on Feature testing and especially on reliability estimation based on Operational Testing. We will create a test procedure which is basically a controller that sets up environmental conditions and invokes randomly selected test cases at random times. We will give higher weight to operations which have higher probability of occurrence. Thus, we will assign a higher number of test cases to them, in order to obtain a good estimation of the platform (server or cloud) performance. The performance we will be closely related to the variable number of operations.

In operational testing a Test Operational Profile must be developed for each operational mode. A Test Operational Profile is different from the Operational Profile for each operation mode due to:

1. The probability of the most common operations
2. Adding new operations

4.3 Test Cases Creation

There exist several criteria from which test cases can be constructed: Design, Code,

Prototype, and Performance.

We will focus on performance test case creation which based on the usage of the software system. In general we will create a certain number of test cases based on the occurrence rate of the operation multiplied with the total number of users performing those operations.

4.4 Application feature testing

The final objective is to give a good estimation of the software reliability. We have considered the feature testing where each operation is going to be tested separately as if it wouldn't influence the others. This is just an assumption which might change on the software deployment case and field testing.

The final objective after the entire testing process is that we can determine if the reliability objective of the entire software has been met. If not further improvement of the software system might be scheduled.

5 Case study and results

In this section we will describe the Use Case diagram of the ACS software system taken as our case study. The Use Case diagram is able to identify the external system functionalities used later on to perform the different reliability analysis. In Fig. 3 is shown the UML diagram constructed with visual paradigm.
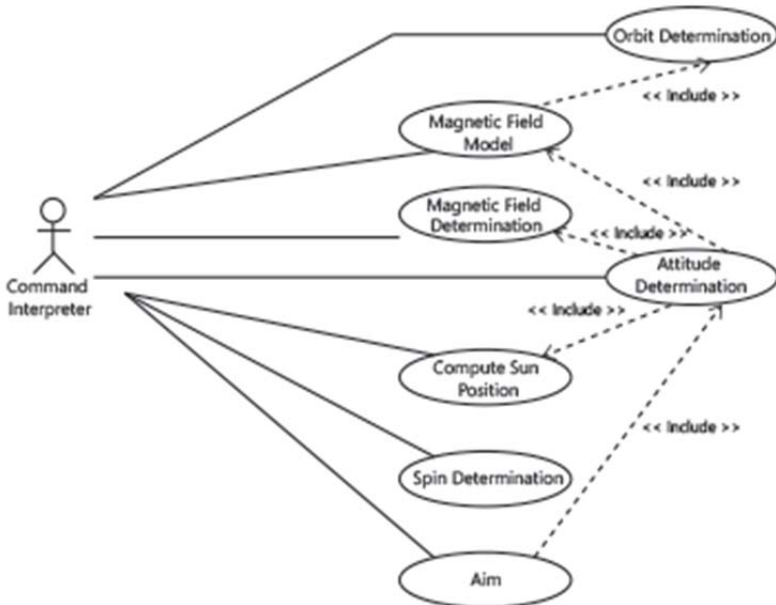
Fig. 3. ACS use case diagram

In the previous diagram we can observe the only actor influencing the system, represented from the command interpreter. The latter is represented from the OBC (On Board Computer) used to decoding the commands arriving from the Earth Station, and transferred throughout the Telecommunication tiles.

In Fig. 4 is given an overview of the classes involved in the construction of the system:
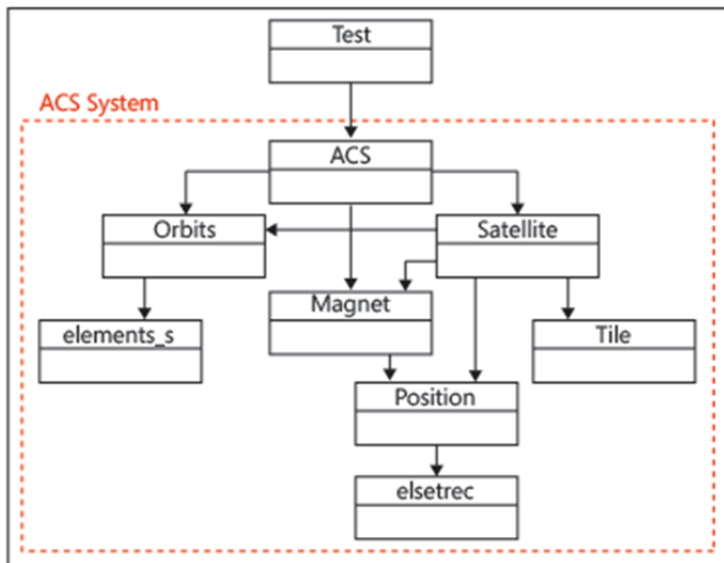
Fig. 4. ACS class diagram

RBD (Reliability Block Diagram) derived from the class diagram and architecture of the software components is shown in Fig. 5. In this case study the RBD has been quite useful in determining the overall expected reliability based on the reliability estimation of each component.
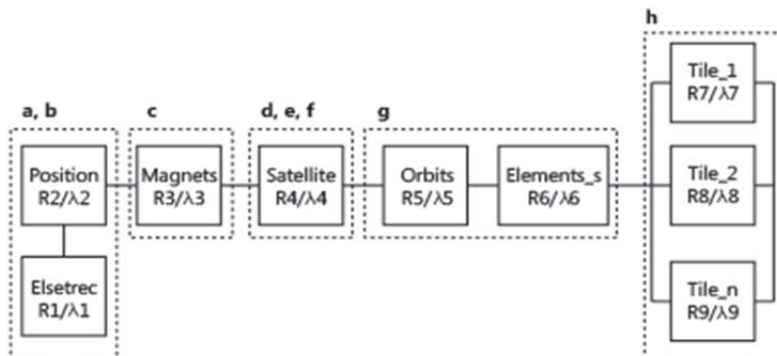
Fig. 5. Reliability Block Diagram of the ACS software system components

5.1 Failure Severity Classes (FSC)

From the RCB (Reliability Block Diagram) and FMEA (Failure Mode Effect Analysis) we can derive the different failure severity classes for the ACS. In Table 1 are reported the possible FSC:

Table 1. Reported the possible FSC

	Failure Classes			
Classification Criteria	**Class 1**	**Class 2**	**Class 3**	**Class 4**
System Capability	F1, F4	F2, F3, F5, F6, F7	F8, F9, 10	
Cost		F1, … ,F10		
Environment			F1, …,F10	
Human Life				F1,…,F10

The identification of the FSC is fundamental to determine the possible reliability requirements of the software system and its failure intensity objective. The overall Reliability of the system is given from:

$$RACS = 0.96$$

taking in consideration that the reliability for each connected block in the RCB is 0.99.
While the FIO (Failure Intensity Objective) according to the formula is:

$$\lambda \approx \frac{1 - R}{t} = 0.00004$$

With such failure intensity objective we can assure that the software is reliable enough for deployment.

5.2 Operational Profile development

The operational profile of the ACS is represented in Table 2. The operational profile was constructed according to the steps described in section 4. The number of operation was determined from code inspection. This operation profile was developed for a normal operational mode of the ACS. It will be further used to guide testing decisions.

Table 2. The Operation Profile of ACS

Operational mode	Initiator	Operational List	Occurrence Rate (Computations per month)	Occurrence probability (Occur. Rate/ Total)
Attitude Computation Mode	On Board Computer	1) Compute Satellite Position	1.30	0.15
		2) Compute Earth Position	96	0.11
		3) Model The Magnetic Field	96	0.11
		4) Compute Magnetic Field	96	0.11
		5) Compute Spin	35	0.04
		6) Compute Day	70	0.08
		7) Compute Sun Position	70	0.08
		8) Compute Attitude	96	0.11
		9) Read Data from sensors	44	0.05
		10) Coordinate the reaction wheels	140	0.16
		Total	**873**	**1**

5.3 ACS Testing and Failure Report

ACS testing for the operational profile constructed above. The total number of test cases is 272 and since this is the first release of the software all test cases are new.
The threshold occurrence probability is:

$$0.5/200 = 0.002$$

There exists no critical operation. The number of infrequent operations with occurrence probabilities below threshold is 0.
The total number of test cases is distributed among the operations based on their occurrence probabilities.
ACS failure report for every test operation according to the previously constructed test operational profile is reported in the Table 3. For the Compute Satellite Position Operation the test documentation is reported in the table 4.

Table 3. Test Case Documents 1

Test Case Document						
Test Case ID: Satellite Position		**Author:** Orges Çiço			**Updated by:** Orges Çiço	
Parent Use Case ID: Computer Satellite Position		**Date:** 1/07/2012			**Updated on:** 12/08/2012	
Test Objective: Determine The Satellite Position Operation						
Unit Nr.	Test Condition	Test Action	Input Specification	Output Specification (Expected Results)	Criteria Pass/ Fail	Comments
1	None	Year Testing	i s.t. 0 <=i <=9	Error message. Computation stopped.	Fail	Computation performed (No check for input errors)
2	None	Year Testing	ij s.t. i>2 0 <=i,j <=9	Error message. Computation stopped.	Fail	Computation performed (No check for input errors)
3	None	Year Testing	ijk s.t. i>2 0 <=j,k <=9	Error message. Computation stopped.	Fail	Computation performed (No check for input errors)
4	None	Year Testing	ijkl s.t. i>2 0 <=j,k,l <=9	Error message. Computation stopped.	Fail	Computation performed (No check for input errors)
5	None	Year Testing	ijkl s.t. i=1 and j=9 0 <=k,l <=9	Sun position Computation within range of degrees.	**Pass**	**Successful**

Table 4. Test Case Documents 2

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
Completed %	17%	33%	50%	67%	83%	91%
Test Cases Passed	50	50	50	50	50	22
Test Cases Failed	0	0	0	0	0	0
Total Test Cases Planned	300	300	300	300	300	300
Test Cases Completed	50	100	150	200	250	272

Since it would be quite exhaustive to display all the test case suite used, the summary graph below, Fig.6 gives an overview of the test case suite construction:
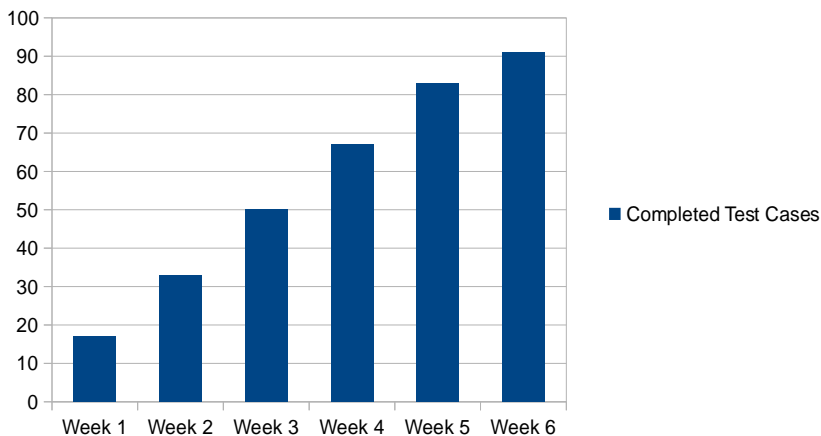
Fig. 6. Completed test cases

Reliability Objective Met?

From the testing reports showed earlier and the testing data collected, the outcome of the experiment proved that the software system had a higher reliability of approximately 0.99 than the reliability objective 0.96 established in the previous section. However the experiment has been conducted on a normal operational mode. Further operational modes must be developed and the overall reliability of the system must be measured, especially for critical operational modes.

6   Conclusions and Future Works

This paper work has enabled the candidate to get acquainted with several concepts concerning Software Reliability Estimation and Testing. Most of the arguments treated are new to the candidate and have not been previously treated during his study. The first part of the paper was dedicated giving a good definition of the concepts regarding system reliability. Most of the concepts have been described so that even a less experienced reader can easily go through understanding the rest of the work. During this part general concepts and definitions of system reliability were explained. Moreover, the exact mathematical measures were explained together with some of the most common probability distribution functions used in reliability calculation. After a good background has been established, the state of the art models and techniques have been reported. This review of the state of the art has permitted to evaluate the different models and techniques developed so far.
Moreover, the associated examples have given a good idea of their effectiveness on real life projects.

Furthermore, reliability testing has been applied on a real case study, according to the methodology explained during this paper work. The case study was the Attitude Control System of a nanosatellite, thus being a system with high reliability requirements in its deployment. The experiment was conducted according to the reliability process proposed from most researchers. The outcome of the experiment was well in the reliability testing. However, the software lacked fault tolerance and did not behave well on some range of erroneous inputs. Thus more robust code needs to be developed. In general operational and reliability testing have yet to prove their effective on real projects. Their cost can be high and the development of the operational profiles might require a lot of effort. The tools developed so far require a considerate number of failures in order to provide with a good Reliability Estimation and the models developed are not totally accurate.

Future work could be made to establish other Operational Modes of the software based on different criteria. This could improve the reliability estimation of the ACS software. Also fault tolerant techniques can be used to observe its reliability increase.

REFERENCES

1. Humphrey, W.S.: The future of software engineering: Watts New Column, News at SEI, 4, March (2001).
2. Lyu, M.R.: Handbook of software reliability engineering. IEEE Computer Society Press and McGraw-Hill, (1996).
3. Tohma, Y.: Models to Estimate the Number of Faults still resident in the Software after Test/Debug Process (1991).
4. Trivedi, K.S., Vaidyanathan, K.., Goseva-Postojanova, K.: Modeling and Analysis of Software Aging and Rejuvenation,. in Proceedings of 33rd Annual Simulation Symposium, IEEE Computer Society Press, Los Alamitos, CA, pp. 270-279, (2000).
5. Haifeng, W.: A software diversity model for embedded safety-critical system Wireless Networks and Information Systems, 2009. WNIS '09. International Conference, pp 106-109, 28-29 Dec. (2009).
6. Xingyuan, Zh.,Jiliang, T., Decun, D.: Software Safety Analysis of 2-out-of-3 Redundant Architecture System Based on Markov Model. Reliability, Maintainability and Safety (ICRMS), 9th International Conference, pp. 493-498, 12-15 June (2011).
7. Scott, R. K.: Investigating version dependence in fault-tolerant software. AGARD 361, (1984).
8. Armoush, A., Salewski, F., Kowalewski, S.: A Hybrid Fault Tolerance Method for Recovery Block with a Weak Acceptance Test IEEE/IFIP International Conference pp 484-491, (2008).
9. Minohara, T., Tohma Y.: Parameter estimation of hyper-geometric distribution software reliability growth model by genetic algorithms. Proceedings of Sixth International Symposium on Software Reliability Engineering, pp. 324 – 329, (1995).
10. Goel, A.L., Okumoto, K.: "Time-dependent error-detection rate model for software reliability and other performance measures,". IEEE Trans. on Reliability, Vo1.R-28, pp.206-211, (1979).
11. Cai, K. Y., Dey, S., Chen, L.: Censored Software Reliability Models PDCS 2:   Open Conference, pp. 555-572, Dept of Computing Science, University of Newcastle, NE1 7RU, UK, (1994).
12. Keiller, P.A., Mazzuchi, T. A.: Enhancing the Predictive Performance of the Goel-Okumoto

Software Reliability Growth Model Reliability and Maintainability Symposium, 2000. Proceedings. Annual, (2000).

13. Lewis, C.A., Smith, R.W., Beaulieu, A.: A model driven framework for N-version programming Systems Conference (SysCon), 2011 IEEE International , (2011).
14. Lyu, M.R.: Software Reliability Engineering: A Roadmap Future of Software Engineering. pp: 153-170, FOSE '07 (2007).

Building Web Based Applications in the Cloud:
A Proposed Model

Case Study: Implementation/Integration of several e-services of SEE University in the Cloud

Agon Memeti[1], Betim Cico[2]

[1, 2] Faculty of Contemporary Sciences and Technologies

South East European University, Tetovo, Macedonia

[1] agon.memeti@seeu.edu.mk, [2] b.cico@seeu.edu.mk

Abstract. This paper presents a very important research problem oriented on building web-based applications in the Cloud, as a quite new discipline and on the focus of many research groups in different institutions and around the world. It is presented during different software's and application implementations and the idea comes for something new, innovative, and contemporary around the world. The aim of this paper is to propose a modelling platform (framework) that is tailored to build efficient, elastic and autonomous applications from tasks and services provided by the Cloud environment, and to define patterns that can result in the efficient optimization of money and resources, because at SEE University there are different web application platforms that serve to offer services to the students but not integrated, they are separated two much in psychical and business process aspects. This will also help in integrating services in a unique framework in the Cloud as a case study of this paper.

Keywords: Cloud Technology, Security, Reliability, Web-Based Applications.

1 Introduction

Cloud Computing is a technology that involves dynamic provisioning of shared computing resources on demand. It is a pay-as-you need model that offers computing resources as a service in an attempt to reduce IT capital and operating expenditures, especially in Universities. The problem is that current software is related to address elasticity, virtualization and billing. To further complicate the problem, each platform provider has different standards that influence the way applications are written. This ties Cloud users to a particular provider. The focus is on defining a proposed model for Cloud Computing, especially implementing/integrating SEEU services (e-learning, administrative and library services) in the Cloud and comparison in performance and cost of services provided by local servers.

The traditional system has several problems such as: lack of mobility, accessibility and portability. These issues can be addressed by the development and use of Cloud

Computing based dynamic software services for SEE University which will facilitate optimum utilization of competent faculty and cover large number of students.

The customized software services will be delivered to encourage and expand the span of e-Learning and other University services (administrative and library services) in the region and around. Using this specific paradigm, Cloud Computing, we will develop the software products on demand with specific requirements and these will be implemented for students from different streams. Our goal is to propose a new model that would support and promote the Cloud based e-Services to improve the present education model preserving the stability of services.

The paper is composed of two sections. The first part is the state of the art of the research paradigm and the next one, the expected outcome of the proposed platform which it will be used for implementation/integration for several SEEU e-services in the Cloud: providing high reusability degree; high performance degree; dynamic integration and increasing portability and interoperability between services looking through different possibilities reducing functional problems by integrating them reusing existing services.

2   State of the Art

Cloud Computing is a technology that allows for greater agility and cost efficiency in management of digital information of any organization or company, through a simple and flexible implementation where services are provided through the "Cloud" (a public telecommunications network) usually the Internet.

It improves the way that many organizations focus on the provision of business through information technology, from computer storage distributed or financial management.

The technology uses the Internet and central remote servers to maintain data and applications. Cloud Computing allows consumers and businesses to use applications without installation and access their personal files at any computer with Internet access [2] , as shown in the Fig.1.
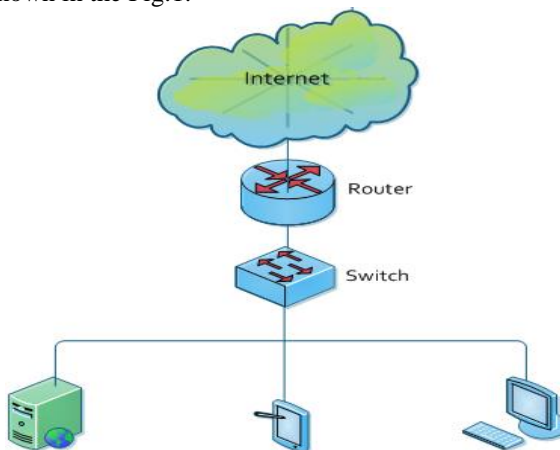
Fig.1. Cloud Computing Scheme

This technology allows for much more efficient computing by centralizing storage, memory, processing, and bandwidth. The technology is not revolutionary, but it is the outcome of the continuous advancement of the data management technology [3].

2.1 Cloud Computing Architecture

Applications built on Cloud Architectures are such that the underlying computing infrastructure is used only when it is needed, draw the necessary resources on-demand, perform a specific job, then relinquish the unneeded resources and often dispose them after the job is done. While in operation, the application scales up or down elastically based on resource needs [4].

These applications run in the Cloud where the provider determines the physical location of the infrastructure. They take advantage of simple API's of Internet accessible services that scale on demand, that are industrial strength, where the complex reliability and scalability logic of the underlying services remains implemented and hidden inside-the-cloud. The usage of resources in Cloud Architecture is as needed, thereby providing the highest utilization with optimum cost [3].

- Everything is a Service
- Separate Compute and Persistence
- Design for Failure
- Design for Resilience
- Operationalize Everything
- Security at Every Layer [5].

2.2 Cloud Computing Models

The services offered by the "Cloud" are distributed among all traditional architectural layers of a computer system, from the hardware layer to application layer software itself. In practice, service providers tend to offer Cloud services that can be grouped into three categories:

- *IaaS (Infrastructure as a Service)* - a model of Cloud Computing, which allows the use of a hardware computing resources as a service provider. Thus, IaaS allows customers to purchase hardware resources (servers, storage, switches, routers, etc) as if it were fully outsourced services. This model has achieved to extend or reduce the physical computing resources a very short period, providing computer infrastructure as a service.

- *PaaS (Platform as a Service)* - is a Cloud-based application development environment. Using a PaaS, we can produce new applications more quickly and with a greater degree of flexibility than with older development platforms tied directly to hardware resources. Running application development on a PaaS has a number of key benefits, starting by the fact that programmers and development managers especially appreciate that the Cloud provider handles all the care and maintenance of the underlying operating systems, servers, storage, and application containers.

- *SaaS (Software as a Service)* - the way of running applications on a Cloud infrastructure, which can be accessible from various client devices such a Web browser. The consumer does not manage or control the underlying Cloud infrastructure, including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

Taking in consideration the fact that Cloud Computing business model is to deliver IT services (software, platform and infrastructure) the potential benefits from Cloud Computing are described below:

- Increased Cost Efficiency
- Increased Provisioning Speed, and
- Scalability

2.3 Types of Clouds

There are several types for the systems that make use of the paradigm of Cloud Computing as shown in Fig.2. The idea is choosing the appropriate model to solve a specific problem.
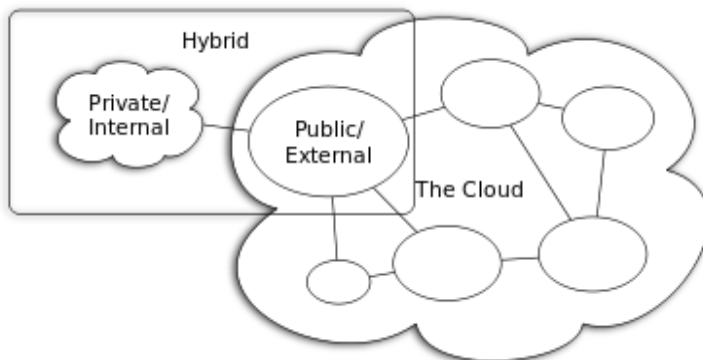
Fig.2. Cloud Computing Platforms [6]

The Private Cloud is established for a specific group or organization and limits access to just that group. Private Clouds are built exclusively for a single enterprise. They aim to address concerns on data security and offer greater control, which is typically lacking in a Public Cloud [7].

The name Public Cloud refers to the standard model of Cloud Computing where services can be available to anyone on the Internet infrastructure (its software or hardware) free or by paying certain amount related to the volume or time of use thereof, while the Hybrid Cloud model is the combination of the two models described above so that the advantage of physical location of the information managed by exploits Private Clouds with the ease of expanding Public Cloud resources.

2.4   Cloud Computing in Universities

Cloud Computing can be used almost in every area. It must be mentioned that (in particular larger) companies already have a functional ICT landscape at their disposal, as a rule. Clouds computing must, therefore, seamlessly integrate it into the existing systems and if not renting storage needed Cloud capacity especially in the implementation of Cloud Computing in University.

The Cloud Computing trend of replacing software traditionally installed on campus computers (and the computers themselves) with applications delivered via the Internet is driven by aims of reducing universities' IT complexity and cost. Cloud Computing could be a technological innovation that both reduces IT costs for the college and eliminates many of the time-related constraints for students, making learning tools accessible for a larger number of students [8]. There are many benefits of Cloud Computing for educational institute and below are listed a few of them;

- With Cloud Computing, universities can open their technology infrastructures to businesses.
- The efficiencies of Cloud Computing can help universities keep pace with ever-growing resource requirements and energy costs.
- The extended reach of Cloud Computing enables institutions to teach students in new, different ways and help them manage projects and massive workloads.
- When students enter the global workforce they will better, understand the value of new technologies.

Cloud Computing allows students and teachers to use applications without installing them on their computers and also allows access to saved files from any computer with an Internet connection [8].

In general the network topology for University labs where several services are used is design in that manner to offer isolated services. The entire buildings are isolated in two logical networks: laboratories subnet and public subnet.

   Main subnet network that supports laboratories environment is the core subnet and is serving to provide base services for laboratories. This isolated subnet is very scalable and it offers redundancy network access for each laboratory in accordance to the service and specifics that each laboratory requires.

   The second logical network it is the public network access subnet. In this topology design resides the public IP subnet of IT department. Using this subnet the IT can publish all services to the Internet and beyond.
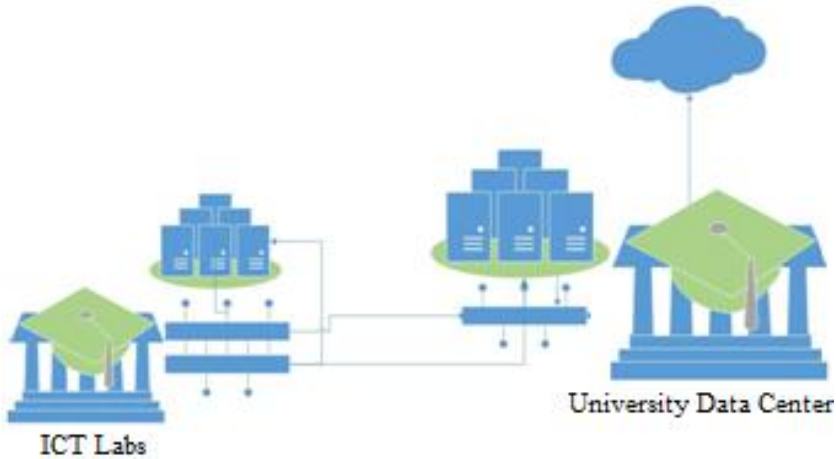
Fig.3. SEEU Data Centre Topology

   As can be seen from the logical design the ICT Lab topology, Fig.3, has to separate links from which is connected to the main University data centre. The first link as described above is the main subnet which connects the laboratories by using distribution layer switch with the core layer switch on the end side University main data centre.

   To fulfil the needs for the laboratories SEE University has implemented a flexible system. This flexible system is running on Sun Microsystems environment technologies hardware and software solution, but this solution offers very flexible administration capabilities since the processing and computing for all users is done in a single platform server.
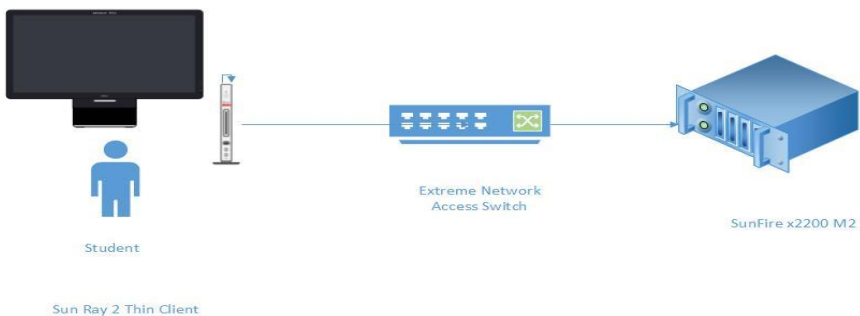
Fig.4. Topology System Design

Fig. 4 shows the technology system design for desktop environment. Students are accessing the system client desktop environment whereas all the processing and computing is done on blade server, where services are installed. Client is connected directly on the access switch and from the switch directly on the server. Thin Clients are used for projecting the desktop view and server.

This versatile virtualized platform in some way offer unified on premise environment where provisioning and automation are the main benefits (for the administration and management part) whereas self-service is what faculty staff and students benefit too.

3 Approach and Expected Outcome

With the rapid evolution of the web and new web technologies such as Web.2.0, AJAX and SaaS, Web based applications are facing new issues such as: Multiplied cost due to the increased in the complexity degree of web applications; as web applications become more and more complex they are increasing in volume, and this implies an increment in the latency degree, making this type of applications useless; web applications require a software infrastructure to public, retrieve and maintain software; security concerns must be taken into consideration to avoid malicious actions and infrastructure is required to provide software availability, we will propose a new model that is tailored to build efficient, elastic and autonomous applications from tasks and services provided by cloud environment by integrating several e-services in the cloud which will preserve stability of the system.

The idea was integrating all services in the cloud in a manner to preserve stability, increase in the portability and interoperability between services and reduction of complexity. As NIST refers portability of across applications, the ability "of prospective cloud computing customers to move their data or applications across multiple cloud environments at low cost and minimal disruption", or the other cloud computing barrier such interoperability between services which according IEEE as "the ability of two or more systems or components to exchange information's/data".

Also the modeling language should be platform independent, with enough technical details that allow it to tackle the platform specific environments. This will facilitate the task of the design and implementation of application in these environments.

Will be used data from various sources, starting from those with an important role in this discipline, as well as those related with this research, such as data and information's from published papers, journals, articles, etc.

The research is conducted on the basis of the approach in Fig.5. Firstly, background on cloud computing and University services will be given and it will help understand the state of the art, such as: approaches and alternatives to achieve the objective; dynamic integration of several services; design and develop a model to support the integration of several e-services in the cloud; evaluate cloud providers and select the best approach and in the end the performance improvements.
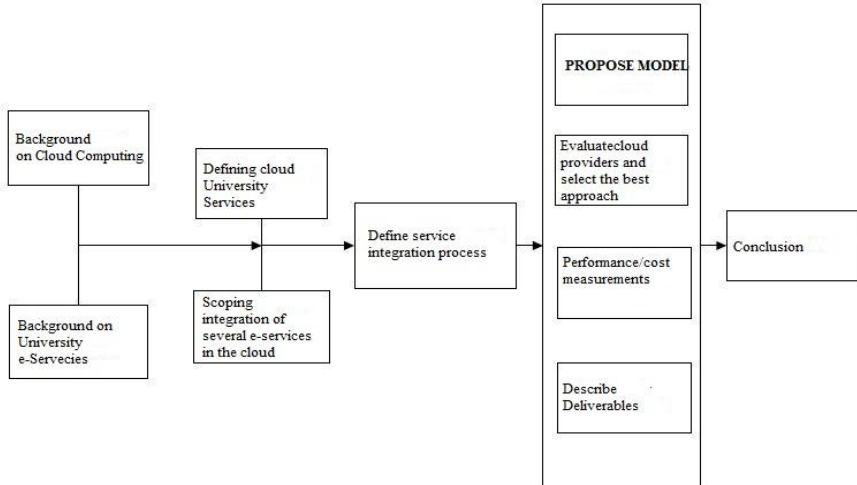
Fig.5. Research Approach

The first aim of our research work is to explore and to design a model from which web-based applications with high consistency; availability and scalability requirements can be developed and deployed on different Cloud infrastructure providers.

Different problems will be addressed, such as: Infrastructure API, by identifying a minimal set of services and corresponding API; application consistency, direct effect on availability but also in performance and cost since we will do a comparison of services provided by normal servers and those integrated in the cloud.
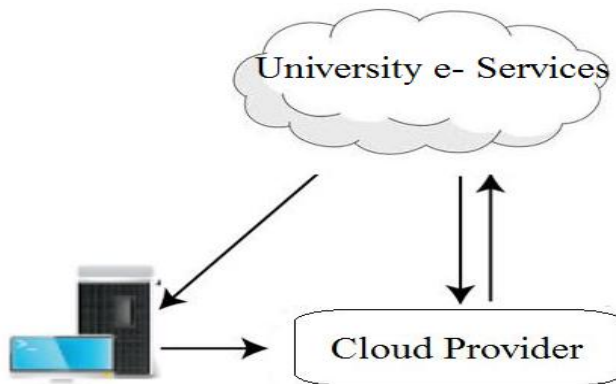
Fig.6. Model of University e-Services in the Cloud

In Fig.6 the user (student, administrative or staff) will send a request to the provider of Cloud Services, than the cloud provider will manage the recourses and services that the user wants and connects to the University e-Service Cloud.

We are committed to deliver the blended integrated e-service solutions for the enhancement of current education model in the region, reducing the cost and development and provide the ability to collect distributed University materials via the internet.

Conclusion

Cloud Computing comes as a great solution for building web based applications and accessing resources in every place as a new approach to produce a solution for old problems.

We have tried to show that the Cloud Computing can also be used for universities and facing web based development process since there are several e-services and they are not integrated.

Using it in educational institutions will promote huge benefits such as high available services, scalability, increasing portability, performance/cost and overall reliability improvements. A few universities already started Cloud Computing technology for educational use. The main goal of suggested model for service integration in the cloud is: managing effectively the functional failures and performance problems of University services.

References

1. Hamdaqa M. et al. A reference Model for developing Cloud Applications. International Conference on Cloud Computing and Services Science. (2011). pp. 98-103.
2. Goyal V. and Kant Ch. Security Issues of Cloud Computing. An International Journal of Engineering Sciences. ISSN: 2229-6913. Issue Sept. (2011), Vol. 4. pp.247-282.
3. Singh H. Current Trends in Cloud Computing – A survey of Cloud Computing Systems. International Journal of Electronics and Computer Science Engineering. ISSN – 2277-1956. pp. 1214 – 1219.
4. Varia J. Cloud architectures. White Paper of Amazon. s3. (2008).
5. Drake J. et al. Developing Cloud-Capable Applications. Open Data Center Alliance. (2012).
6. http://en.wikipedia.org/wiki/Cloud_computing
7. Harris T. Cloud Computing – An Overview. SOA. http://thbslb-2119610767.us-east-1.elb.amazonaws.com/knowledge-zone/cloud-computing-overview
8. Erkoc M. et al. Cloud Computing for Distributed University Campus: A prototype Suggestion. International conference – The Future of Education. http://conference.pixel-online.net/edu_future/common/download/Paper_pdf/ENT30-Erkoc.pdf

Storage Based Cloud Computing and Disaster Recovery

Artur Koçi[1], Betim Çiço[2]

[1,2] Computer Engineering, Epoka University,Tirana, Albania.
[1]akoci@epoka.edu.al, [2]bcico@epoka.edu.al.

Abstract. Cloud Storage as a part of cloud computers based platform is used to solve many problems of business services, such as Storage capacity, security, load, performance and many other issues. Business data services are migrated into Cloud and powerful operating systems and platforms are used to process the data. This paper gives an overview of cloud storage-based platforms and presents a new proposal model for disaster recovery and other applications in inter collaboration private cloud.

Keywords: Cloud computing; cloud storage;disaster recovery; inter collaboration private cloud*.*

1 Introduction

Cloud computing is an emerging model for business computing [1]. The task to be computed is distributed in a large number of computers in the cloud, where all applications can access the storage space and software services.

Cloud computing systems provide a powerful and reliable network service similar to a super computer solution. Cloud computing usually is based on the proceeding of the user`s request and, it gives the output result.

Cloud storage is a new concept extended and developed from the concept of cloud computing [2]. Cloud storage has a large number of storage devices which are configured together and a large number of data are stored on it. Cloud storage is part of cloud computing which is interested only in data storage and their management. Cloud storage system itself has to include network equipment, storage equipment, server applications, and public access interface and client programs.

Cloud storage system gives services to end users, to store local data in the online storage space provided by storage service provider through network.

In this research it is given an overview of cloud storage-based platforms architecture, and also a new model is offered as a solution for disaster recovery implementation. The model offered in this research is more economic, and can be applied in companies with limited budgets.

2 Cloud Storage System

Cloud Storage system uses software to provide interconnections and collaboration between different types of storages. For making different storage working and collaborating together, to provide data storage and business functions, cloud storage uses application software, thus, bringing a new challenge and a new idea regarding data security, reliability and data management.

Implementing cloud storage system, the details of each storage device should be known. The detail of storage device consists on: the model of the storage we are using, the interface being used and, the transmission protocols, the number of disks, and the full capacity, finally on the kind of connections cable it uses between storage and server etc.

From the user side using cloud storage, as disaster recovery solution, it is no need to build their data backup and take care of the integrity of the data. No need to ensure data security and connection availability. Users also do not need to preoccupy for device equipment condition monitoring, software upgrades etc. All the devices in the cloud storage are transparent and any user that has authorization can connect to the cloud storage and access data on the storage in any place.

Cloud storage system architecture model consists of four logical layers, as shown in Fig.1.
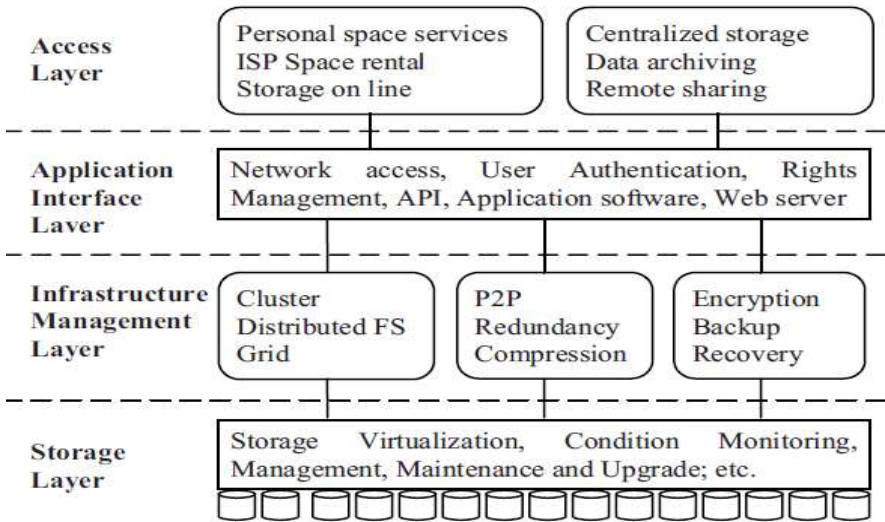
Fig. 1. The architecture model of cloud storage system [3].

1- Storage layer - The basic layer of cloud storage system is the storage layer. Storage devices connected to the cloud storage system can be fiber channel (FC) Storage devices, IP devices such as iSCSI, and NAS. Moreover, there can also be DAS storage devices such as SCSI or SAS. Cloud storage system consists in a large number of storage devices geographically distributed. All the devices connected with each other through the wide area network (WAN). In the Storage layer is the unified

storage device management system, which can realize the virtual storage management, multi-link redundancy management, hardware equipment condition monitoring, and fault maintenance.

2- Infrastructure Management Layer - The core part of the cloud storage system is Infrastructure Management Layer this layer uses clusters, distributed file system and grid computing technology to archive cooperation between multiple storage devices in the cloud storage and also to provide data access high performance. Insurance of the data from the access of the unauthorized users is provided by content distribution system, and data encryption technology. Data back up and disaster recovery measures can guaranty that the data stored in the cloud storage will not be lost.

3- Application Interface Layer - Application Interface Layer is the flexible interface layer of the cloud storage. Application interface layer can be different for different storage provider and mostly the application service differs based on the business type.

4- Access layer - Authorized user can login the cloud storage system via standard application interface to get cloud storage services. Different cloud storage provider provides different access types and access methods. The main function of cloud storage is to share data. Common file system only allows visiting data exclusively, but commercial applications require sharing data between the operating system and the "data warehouse". Different copies of data on different server with different operating system should overcome the I/O access conflicts when data were share and stored in parallel memory. This requires a good lock algorithm, multi-level locking mechanism and Cache consistency techniques to ensure the consistency of the data.

3 Disaster Recovery Model of Cloud Storage

To ensure the data security and integrity on cloud storage system is very difficult to. Because of the transparency of the applications and data services, the malicious attacks must be avoided. Cloud storage system must have its own fault tolerance which is necessary to be with a low cost.

At the same time, there must be Fault tolerance backup system to avoid the data loss because of the force majeure [1], [4].

To ensure the continuity of application and the security of data, the structure of disaster recovery system is "distributed computing, centralized storage". Disaster recovery has three levels such as data-level disaster recovery, system level disaster recovery, and application-level disaster recovery [5]. A typical architecture of disaster recovery system is shown in Fig. 2.

The architecture uses Storage area network (SAN) to support the application data access and backup. SAN system uses high performance routing devices and redundant links for application server to access the file server and backup data. SAN as a large, stable data transmission network requires a number of high-performance routing equipment to build and the hardware cost is relatively high. From a technical development trend, the cloud storage architecture has high reliability, high performance, which are all easy to expand. The online backup service based on it can

adapt to Web 2.0 requirements. The online backup service cloud-storage-based can provide dynamic scheduling, standardized access interface, and dynamic memory map management [3].
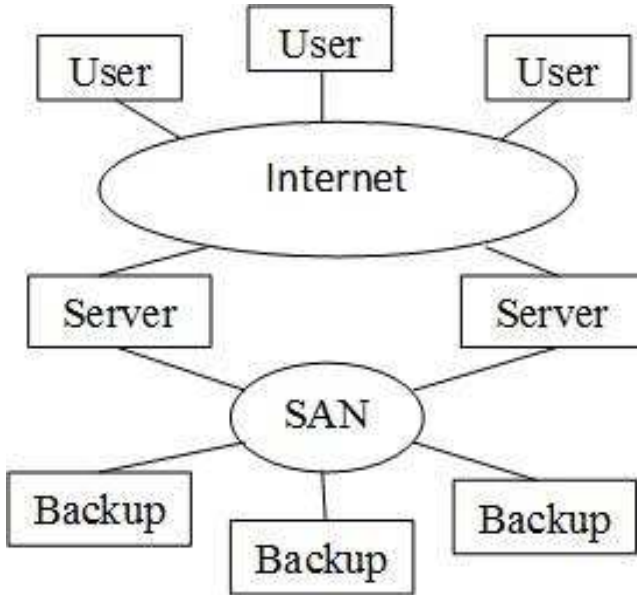
Fig. 2. The typical architecture of disaster recovery system.

3.1 Disaster Recovery Models

Many authors have studied and proposed different solutions for disaster recovery in cloud computing, the field of study is very large and differs based on the purpose of cloud usage.

In their research paper, Pokharel, M., Seulki Lee, Jong Sou Park are offering two zones, Zone A and Zone B [6]. Each zone is replication of another. If one zone gets down due to disaster then another will be up. Monitoring Unit: In order to provide more reliable disaster recovery system, they propose another unit known as monitoring unit that monitors the functionalities of both zone "A" and zone "B". It takes the response from load balancer about the resource constraints. Once it gets the request from load balancer, then it verifies it again aiming to see the need for load balancing. Afterwards, it places the order back to load balancer to execute accordingly. It also monitors the state of both zones this means it monitors the functionalities of system architecture (SA). If it finds there is disaster at zone "A" or not in good state or it is in compromise state then it does make zone "B" ready to take the charge. So, before zone "A" is completely in failure state, zone "B" becomes active and provides the services to the citizen [6].

In another research paper, Zhang Jian-hua and Zhang Nan [3] are offering the model that consists on building inter-private cloud storage as in Fig. 3.
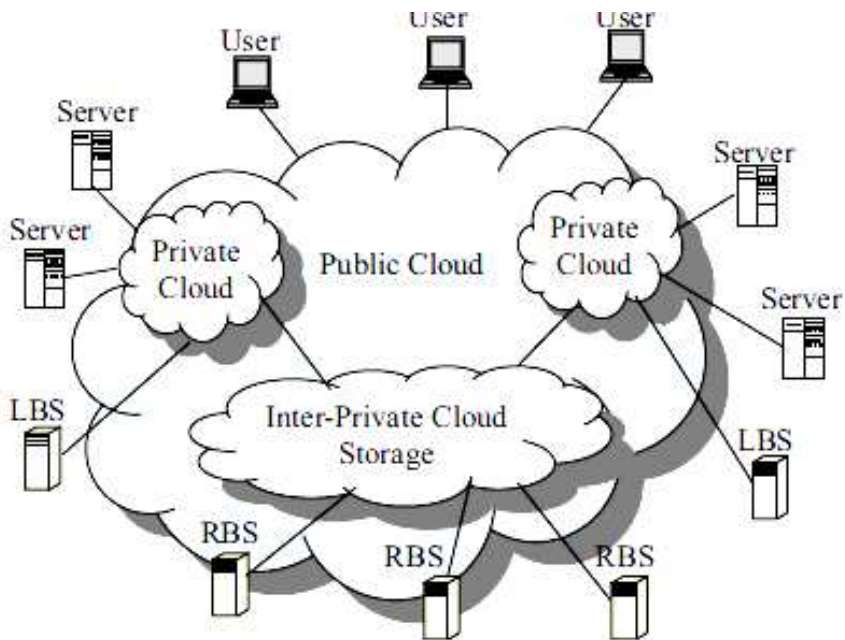
Fig. 3. The model of disaster recovery system [3].

The entire servers in the private cloud computing system are equipped with Local Backup Server (LBS) and the Remote Backup Server (RBS).

Services will be borrowed from the inter-private cloud storage. Inter-private cloud storage is cloud storage system that offers as a service online storage and fulfils the requirement of having three remote backup servers geographically distributed [3].

3.2 Proposed Disaster Recovery Model

Based on SNIA Standards (Storage network industry association), Disaster recovery must consist on at least three different Remote Backup Server distributed in different geographical location. If private cloud storage is equipped with all the disaster recovery solution and redundant backup system, it will have a very high cost, heavy configuration, resources, lack of flexibility etc.

In this paper we offer a model that consists on gathering together private cloud computers as in Fig. 4, and share cloud storage services between each other. Several private cloud computers, a minimum four private cloud computers agree together to share their services and create a public cloud computing. In the proposed model, Local Backup Server (LBS) of each one of the private cloud computers is used as Remote Backup Server (RBS) for the three other remaining private cloud computers and so on. In the Fig. 4 only one private cloud computer is presented with all the connection.
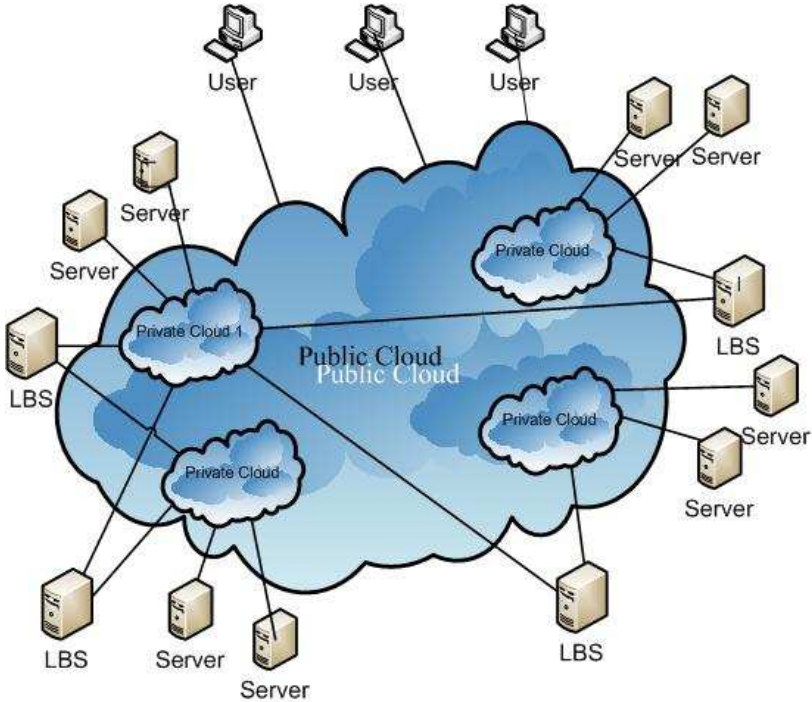
Fig. 4. The Model of Disaster Recovery.

To make the figure more understandable, the other connection between private cloud computers are not shown in the model of disaster recovery.

Application data of the system are stored in server, all servers are equipped with local backup server and three remote backup servers, remote backup server are geographically distributed and the collaboration between private cloud computers is selected in a way that cloud computers shouldn't be in the same geographical area

Data are stored in the local backup server and the data integration to the remote backup server is done in a leisure time, including data delete, disk compression and finishing work

After the initial full back up, incremental backup can be used and synchronization with the remote backup server can be established. Synchronization can be live synchronization or in a specific time where there is less bandwidth usage. In the hardware architecture, connection between private cloud computers does not require high-speed data processing in real time; therefore, it is no need for very high bandwidth connection.

To protect data during transmission all data and transmission channels are encrypted and also different policies are applied to ensure data security.

The Model of Disaster Recovery proposed in this paper is a sustainable solution for private cloud computing with limited budget. Using the proposed model, private cloud computers will eliminate the need of obtaining services from cloud storage for implementing disaster recovery solution. The second one: implementing the proposed model will make possible the efficient usage of cloud computing resources.

4 Conclusions

Nowadays for businesses with limed resources, cloud storage seems to be a good solution, but still there is a risk of losing data, and the implementation of disaster recovery becomes a must As the budged of the companies are limited and also human resources are limited new models should be developed and implemented. The model we offer in this paper can be with high profit for the companies with limited budges because it eliminates the necessity of hiring remote backup servers but still the cloud storage is equipped with disaster recovery solution

As the data in the remote backup are been transferred from the secure channels, the security of the data is increased.

The given model needs Local backup storage to be equipped with more storage capacity

References

1. Armbrust, M., Fox, A., Griffith, R.: Above the Clouds: A Berkeley View of Cloud Computing. University of California Berkeley, CA,USA. (2009)
2. Zhou Ke, Wang Hua, Li Chunhua.: Cloud Storage Technology. ZTE Communications. (2010)
3. Zhang Jian-hua Zhang Nan.: Cloud Computing-based Data Storage and Disaster Recovery. International Conference on Future Computer Science and Education. (2011)
4. Qu Ming-cheng, Wu Xiang-hu, Liao Ming-hong.: A Disaster- Tolerant Storage Model and a Low Data Failure Model for Data Grid. Acta Electronica Sinica. 38(2), pp 315-320 (2010)
5. Bajpai, A., Rana, P., Maitrey, S.: Remote Mirroring: a Disaster Recovery. International Journal Of Advance Research In Science And Engineering. (2013)
6. Pokharel, M., Seulki Lee, Jong Sou Park.: Disaster Recovery for System Architecture using Cloud Computing. 10th Annual International Symposium on Applications and the Internet. (2010)

Big Data Solutions for Cloud Application Platforms

Rustem Dautov[1], Iraklis Paraskakis[1], and Mike Stannett[2]

[1] South East European Research Centre (SEERC),
CITY College – The University of Sheffield International Faculty
24 Proxenou Koromila, 54622 Thessaloniki, Greece
`rdautov@seerc.org`, `iparaskakis@seerc.org`
[2] University of Sheffield, Department of Computer Science,
Regent Court, 211 Portobello, Sheffield S1 4DP, UK
`m.stannett@sheffield.ac.uk`

Abstract. With data volumes increasing exponentially in recent years and existing technologies becoming overwhelmed by the rise of Big Data, cloud computing has come to be seen as a convenient environment for running data– and compute-intensive applications. However, their continuing growth in size and complexity means that cloud platforms are rapidly reaching the stage where they not only host such applications, but also act as generators of Big Data in their own right, continually emitting terabytes of raw heterogeneous data at unpredictable rates. As a result, novel approaches to the monitoring and analysis of cloud platform behaviours are required. Taking Heroku as an example (a cloud application platform with over one hundred built-in and third-party services and more than a million deployed applications) we consider this situation in detail, and outline a possible framework for resolving the problem. The presented framework employs an existing Big Data technology to process large amounts of monitored data streams within Heroku in a parallel fashion. As data volumes increase and performance may drop, this technology combination has the potential to address the scalability issue and facilitate timely reactions to critical situations.

Keywords: Cloud Computing, Cloud Application Platform (CAP), Heroku, Big Data.

1 Introduction

The past decade has seen explosive growth in Internet traffic volumes (Fig. 1), with companies, institutions and households exploiting the ever-increasing variety of resources and services available online. Global mobile traffic alone is expected to rise from 1.5 exabytes $(1.5\,\mathrm{EB})$[3] per month in 2013 to $15.9\,\mathrm{EB}$ by 2018.[4]

---

[3] 1 exabyte $(1\,\mathrm{EB})$ = 1000 petabytes $(1000\,\mathrm{PB})$ = 10^{18} bytes.

[4] `http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html`

In an effort to handle the associated complexities, providers have increasingly adopted a service-oriented approach, leading ultimately to today's focus on cloud computing. By exploiting virtualisation techniques and economies of scale, cloud computing provides customers with seemingly infinite computing capabilities, while avoiding the costs of acquiring dedicated hardware and the associated large human integration expenses. These features continue to encourage enterprises, governments and research institutions to migrate their IT systems to the cloud; it is predicted that by 2017 nearly half of large enterprises will have cloud deployments.[5]
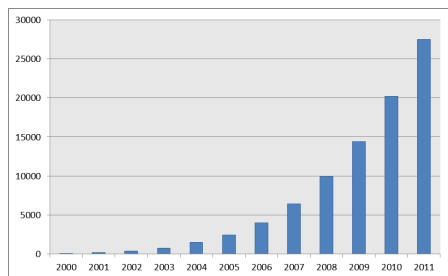
Fig. 1. Global Internet Traffic, 2000-2011 (PB per month). Data: Cisco Systems.[7]

This increasing reliance on cloud-based services and applications emphasises the need for reliability at all levels of the cloud service model, and this in turn highlights the challenge of ensuring accurate and timely data monitoring and analysis. This is particularly important for Cloud Application Platforms (CAPs) [6], where the principles of the service-oriented computing paradigm (the provision of discoverable, reusable and highly interoperable services as basic building blocks from which to assemble rapid and low-cost, yet secure and reliable, applications [10, 14]) are reflected in the large number of generic services available for re-use and integration into users' apps. These include such services as *data storage*, *caching* and *back up*, *searching*, *E-mail* and *SMS*, *logging* and *queue messaging*, among others. In addition, given the wide variety of user requirements, cloud providers typically make APIs and client libraries available in multiple programming languages.

As a result, the integration of services can generate complex interrelationships between cloud platform services and user applications. For example, Google App Engine[8] currently offers 41 services (or "features"); IBM Bluemix[9] provides 34 built-in, 10 community and 15 third-party services; and Heroku[10] offers up to 150 add-on services. The inherent problems are compounded still further by the availability of *hybrid* clouds, where different parts of a cloud-based application system are deployed on private and public clouds, and *multi-clouds*, where a cloud-based application system is distributed across several clouds [13, 5].

In this context, the challenge of maintaining ever-expanding CAP software environments dictates the need for platform providers to maintain constant vigilance over all critical activities taking place on the platform, where necessary introducing new services and applications and modifying existing ones [7]. Given

---

[5] http://www.gartner.com/newsroom/id/2613015
[8] http://appengine.google.com/
[9] https://ace.ng.bluemix.net/
[10] http://www.heroku.com/

that meeting these challenges will increasingly involve the rapid processing of extremely large quantities of real-time data, we can also regard the monitoring and analysis of service-based cloud environments as an exercise in 'Big Data' information processing.

2 Cloud Application Platforms as Generators of Big Data

The term 'Big Data' was first introduced in 1997 by Cox and Ellsworth [2], in the context of datasets around 100 GB in size – at the time, this was large enough to make storage and processing of the data beyond the scope of standard contemporary approaches. Since then, the amount of data that can be handled easily has grown by a factor of around ten thousand, but the explosion in data generation has far out-stripped even this impressive growth rate. Manyika et al. [9] estimate that around 14.2 EB of new data were stored by enterprises and consumers in 2010, and it has been predicted that total annual Internet traffic will have reached 667 EB in 2013 [3]. IBM[11] estimates that the world creates 2.5 EB of data every day, which results in urgent new research questions as to how such extreme volumes of data can be stored and processed in a reasonable amount of time.

The key factors characterising Big Data, and affecting how it should be processed, have become known as the *four Vs*, namely:

– VOLUME – the scale of the data is large, typically in excess of 1 EB;
– VELOCITY – analysis of rapidly streaming data is required in realtime;
– VARIETY – data is stored in many different forms and formats;
– VERACITY – data is often uncertain, flawed, or rapidly changing.

These factors, so central to the concerns of Big Data, are also encountered when dealing with data management and analysis on cloud application platforms, as can be seen by considering each of the four Vs in turn [4].

– VOLUME: *PageLever*,[12] a Heroku-based analytics platform for measuring a brand's presence on Facebook, processes 500 million Facebook API requests each month, which are then stored in a database. Heroku itself hosts more than one million deployed smaller-scale applications[13,14] and offers more than one hundred add-ons. Even if we neglect noise, i.e. data flows which are not relevant for monitoring in the given context, the amount of data remaining is still considerable.
– VELOCITY: *Quiz Creator*,[15] another application deployed on Heroku, reports activity peaks of over 10,000 user requests/minute. Similarly, *Playtomic*[16]

---

[11] http://www.ibmbigdatahub.com/infographic/four-vs-big-data
[12] http://success.heroku.com/pagelever
[13] http://gigaom.com/2012/05/04/heroku-boss-1-5m-apps-many-not-in-ruby/
[14] http://www.crunchbase.com/company/heroku
[15] http://success.heroku.com/cardinal_blue
[16] http://success.heroku.com/playtomic

claims around 15-20 million gamers generating over a billion events daily. As these examples demonstrate, CAPs must handle multiple rapidly-generated streams of data (which are then processed, stored, deleted, etc.) at an unpredictable rate. Moreover, platform components are constantly evolving and changing, new services are added and old ones are removed, new sources of raw data constantly emerge, and old ones are disappearing, thus making the whole system even more dynamic.

– VARIETY: the data originating from the various sources on a CAP (applications, databases, users, services, etc.) can be extremely heterogeneous, both in terms of representation (i.e. data formats and encodings) and in terms of semantics (e.g. two applications may store data in the same database – that is, adopting the same format and structure – but belong to completely separate business domains).

– VERACITY: CAPs are time-critical environments in which it is crucial to distinguish between data which is still valid, accurate and actual, and data which is already obsolete, since reacting to out-dated observations may lead to unnecessary (or even harmful) adaptations. It is also important to react correctly to the presence of "noise" so as to avoid any misinterpretation of monitored data.

Of course, the four Vs are not enough, in themselves, to characterise the problems of CAP management. For example, cloud-based data sources can also be characterised in terms of their *distribution*: data may come from various logically and physically distributed sources. The data needed to ensure reliable CAP management may originate from, e.g., databases, file systems, running applications and external services, and these "logical" sources may be deployed on separate virtual machines or web servers, or even reside in different data centres and jurisdictions.

3   Autonomic Management in the Cloud

We have argued extensively elsewhere that the growing complexity of ensuring CAP reliability necessitates the introduction of a more *autonomic* approach to CAP management, and have recently conducted various supporting proof-of-concept experiments to validate our approach [4–6]. While these small-scale experiments have been successful, the challenge of demonstrating scalability inevitably brings with it the challenge of coping with extremely large volumes of data continuously generated and consumed both within and outside the platform. In this respect, the relationship between CAP management and Big Data is highly beneficial, because the Big Data community has already generated efficient solutions addressing the processing and analytics of static data (e.g., Apache Hadoop[17]), and there are several emerging approaches for dealing with streaming data (e.g., IBM InfoSphere Streams[18], Apache Storm[19]). We believe

---

[17] http://hadoop.apache.org/

[18] http://www.ibm.com/developerworks/bigdata/streams/

[19] http://storm.incubator.apache.org/

these solutions to be applicable to the domain of cloud platform monitoring, allowing us to benefit from these highly-optimised and reliable technologies. Our experimental validation of this claim focuses primarily on Heroku as a representative example for experimental purposes, but it should be noted that other CAPs face similar problems [1].

The exploitation of Big Data solutions to support autonomicity in CAPs in our experimental validation is two-fold. First, a fundamental underpinning of our approach, which enabled us to apply an existing Big Data solution from the research area of Sensor Web, is our interpretation of CAPs as distributed networks of *software sensors* services, deployed applications, platform components, etc., are treated as continual generators of raw heterogeneous data which has to be monitored and analysed to support run-time situation assessment. This solution relies on homogenising and providing enhanced meaning for streaming sensor observations by annotating them with Semantic Web languages (i.e., OWL and RDF). This in turn allows us to: (a) perform run-time *stream reasoning* [12] over these semantically-enriched data streams so as to detect critical situations; and (b) diagnose the problem and suggest potential adaptation actions, by reasoning over a set of corresponding SWRL rules.[20]

The initial experiments showed that with the Stream Reasoning approach we were able to detect, diagnose and suggest a possible adaptation action to a critical situation taking place within the CAP within 1 second. However, with the increase in the number of incoming RDF triples the performance of the framework decreases, and at the rate of 1000 triples/second the autonomic framework no longer could perform within the frame of 1 second. Existing experiments suggest that with the increase in RDF data sets from 10K to 1M triples, the average execution time of continuous SPARQL queries increases at least 50 times [8]. Such performance drops make our framework potentially incapable of monitoring and analysing large data sets within CAPs and require an additional solution to be applied to address the scalability issue.

A possible solution to overcome the above-mentioned problems is to parallelise reasoning tasks across several instantiations of the autonomic framework [11] by fragmenting incoming data streams into sub-streams, so that each instance only deals with a separate subset of incoming values. Unlike static data fragmentation, where the set of values is finite, partitioning of streamed data, due to its unbounded nature and unpredictable rate, is associated with a risk of splitting semantically connected RDF triples into separate streams, which in turn may result in incorrect deductions. Therefore, careful design of the fragmentation logic is crucial in order to confirm that no valuable data is misplaced or lost.

Taken together, these considerations suggest that we need to employ an existing solution which exhibits following characteristics:

- Minimal effort to integrate with our framework;
- Support for processing streamed data;

---

[20] For a more detailed overview of our approach we refer the reader to [5, 6].

– Support for data stream fragmentation and task parallelisation;
– Enough capacity to address the Big Data challenges of CAPs;
– Mature level of functionality, support and documentation;
– Free access for research and academic purposes.

By surveying existing technologies, we concluded that IBM InfoSphere Streams (henceforth Streams) meets all these requirements. It is a software platform that enables the development and execution of applications that process continuously flowing data. It supports dynamic analysis of massive volumes of incoming data to improve the speed of business insight and decision making. Streams consists of a programming language (Streams Programming Language - SPL), an integrated development environment (IDE), and a run-time environment that can execute SPL applications in stand-alone or distributed modes. The Streams Studio IDE includes tools for creating visual representations of applications, consisting of operators and connecting streams, by means of a drag-and-drop mechanism. Alternatively, users can also develop Streams applications directly in SPL. What really makes Streams different from other approaches and made us continue further experiments with it, are:

– A wide selection of pre-compiled operators, ranging from simple utility operators (e.g., source, sink, filter, aggregate, split) to more complex ones, such as XML Parser or operators dealing with database access.
– Support for custom operators, which can be written in SPL, Java or C++. Thanks to this feature it is possible to package the autonomic framework as a separate JAR file and access it as an operator from within a Streams application.
– Support for scalability: provisioning of an additional analysis component is a just a matter of dragging and dropping it on the working area - the background routine of configuring and managing of computational nodes is done by the Streams platform, whereas in open-source solutions, this process (known as *zoo keeping*) has to be done manually.
– Support for task parallelisation: with various built-in operators responsible for splitting, filtering and merging of incoming data streams it is possible to achieve highly parallelised behaviour of a Streams application.

As a result, by leveraging Streams we are able to fragment incoming RDF streams and process them in parallel, thus addressing the scalability issue associated with formal reasoning. By splitting the incoming stream of RDF triples into three separate sub-streams and processing them in parallel, we again managed to keep the "reaction time" of the framework within 1 second. Moreover, IBM offers a free Linux distribution with the pre-installed Streams environment for non-commercial use, which has exactly the same set of features as the full-blown commercial offering with the only difference that it runs on a single virtual machine, rather than on a distributed cluster of physical nodes or in a cloud. Therefore, with certain assumptions we can claim that our local experiments can be reproduced at a larger scale if needed. The fact that Streams has been developed by IBM - one of the Big Data pioneers - also gives us confidence that it is a worthy direction to follow.

4   Summary and Conclusions

In this paper we have argued that the continued robustness of CAPs requires the equipping of platforms with autonomous management capabilities, and this in turn requires the proper harnessing of Big Data principles. The presented considerations suggest strongly that data monitoring and analysis in CAPs demonstrate the characteristics associated with Big Data processing, and can therefore benefit from the highly-optimised and reliable technologies developed in this field.

As one of such promising technology combinations, we presented our autonomic framework, which performs run-time analysis of data monitored within the cloud environment of Heroku, and is deployed on IBM Streams – a software platform for processing continuously flowing large amounts of data in a distributed and parallel manner. We demonstrated how the application of Streams could help us keep the "reaction time" needed to detect a potential problem within 1 second by splitting ever-increasing volumes of incoming monitored data into sub-streams and analysing them in parallel.

Acknowledgements

References

1. Brazier, F., Kephart, J., Van Dyke Parunak, H., Huhns, M.: Agents and Service-Oriented Computing for Autonomic Computing: A Research Agenda. IEEE Internet Computing 13(3), 82–87 (2009)
2. Cox, M., Ellsworth, D.: Application-Controlled Demand Paging for Out-of-Core Visualization. In: Proc. Visualization '97. pp. 235–244 (1997)
3. Data, data everywhere: a special report on managing information. Analyst report, The Economist (27 Feb 2010), `http://uk.emc.com/collateral/analyst-reports/ar-the-economist-data-data-everywhere.pdf`
4. Dautov, R., Paraskakis, I., Kourtesis, D., Stannett, M.: Addressing Self-Management in Cloud Platforms: a Semantic Sensor Web Approach. In: Proceedings of the International Workshop on Hot Topics in Cloud Services (HotTopiCS'13), April 20-21, 2013, Prague, Czech Republic. ACM (2013)
5. Dautov, R., Paraskakis, I., Stannett, M.: Towards a Framework for Monitoring Cloud Application Platforms as Sensor Networks. Cluster Computing Journal (2014), in press.
6. Dautov, R., Paraskakis, I., Stannett, M.: Utilising stream reasoning techniques to underpin an autonomous framework for cloud application platforms. Journal of Cloud Computing (2014), in press.
7. Kourtesis, D.: Towards an ontology-driven governance framework for cloud application platforms. Technical Report CS-11-11, Department of Computer Science, The University of Sheffield (2011)

8. Le-Phuoc, D., Dao-Tran, M., Parreira, J.X., Hauswirth, M.: A native and adaptive approach for unified processing of linked streams and linked data. In: The Semantic Web–ISWC 2011, pp. 370–388. Springer (2011)

9. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A.H.: Big Data: The next frontier for innovation, competition, and productivity. Technical report, McKinsey Global Institute (2011)

10. Papazoglou, M.P., Traverso, P., Dustdar, S., Leymann, F.: Service-oriented computing: a research roadmap. International Journal of Cooperative Information Systems 17(2), 223–255 (2008)

11. Urbani, J.: Scalable and parallel reasoning in the semantic web. In: The Semantic Web: Research and Applications, pp. 488–492. Springer (2010)

12. Valle, E.D., Ceri, S., Barbieri, D.F., Braga, D., Campi, A.: A first step towards stream reasoning. In: Domingue, J., Fensel, D., Traverso, P. (eds.) FIS. Lecture Notes in Computer Science, vol. 5468, pp. 72–81. Springer (2008)

13. Vičič, J., Brodnik, A.: Hybrid cloud monitoring. In: Proceedings of the 1th International Conference on CLoud Assisted ServiceS (CLASS), October 22-25, 2012, Bled, Slovenia (2012)

14. Wei, Y., Blake, M.: Service-Oriented Computing and Cloud Computing: Challenges and Opportunities. IEEE Internet Computing 14(6), 72–75 (2010)

Use of ICT in learning contexts: interaction between learning traits and attitudes towards VLE

Christian Beyle

Department of Psychology, University of Sheffield
Western Bank, Sheffield S10 2TN, UK
cabeyle1@sheffield.ac.uk

Abstract. The interaction between learner attitudes towards ICT and cognitive characteristics is a relevant issue when designing effective virtual learning environments (VLE) with the aim of get high rates of adoption. The present research attempts to test the idea that both "learning approaches" and "attitudes towards ICT" interact to determine the intention of adopting a VLE.

228 participants took part in this online study completing a measure of "learning approach" (two dimensions, deep and surface approach) before being randomized to receive either a structured or unstructured VLE in which they faced a hypothetical task. Subjects then reported their attitudes toward the assigned setting.

Results suggest that VLE design has a significant effect on all attitudes towards VLE adoption. Nonetheless, the effect of VLE design is different when the learning approach of the user is considered. Participants with a "deep approach" have better global attitudes towards VLE adoption and better intention of use it than those with "surface approach". The theoretical and practical implications of conceiving a learning technology adoption model that includes an interplay between learner and the virtual environment are presented.

Keywords: Virtual Learning Environment, Attitudes towards ICT, Learning, Adoption of technology.

1 Introduction

Some of the main advantages of information and communication technologies (ICT) with learning purposes rest on its design flexibility, ease of distribution and its potential decrease of costs [1]. These reasons – among others – support its deep introduction over the last two decades into schools, universities and other organizations in order to conduct or complement instructional processes in the form of Virtual Learning Environments (VLE) or Learning Management Systems (LMS) depending on their characteristics and aims. Nonetheless, concerns about its effectiveness and capacity of being suitable for all kind of learners have been set more than once [2].

In this context, two main challenges arise for computer-supported instruction. One of them is to improve the rate of adoption and continued use of VLE, or said in other

words, to decrease the rate of dropouts, since the first requirement for an effective learning process is that people take part on it. This is of special interest for the design of Massive Open Online Courses (MOOCs), which are considered as a revolutionary education scheme, but which dropout rate can be higher than 90% [3], [4].

The other one is to enhance its effectiveness, allowing learners not only to attain the declared learning goals, but also to achieve a significant learning experience. This is, to go further the mere fulfilling of requirements of task and sequences and to engage students into a process of development of knowledge according to their expectancies and goals, assuming that not all students have the same learning styles and objectives.

There is an important number of researches and theory development on VLE design investigating how specific setting configurations affect the way the learners achieve instructional goals [5]–[10]. The assessment of practitioners' orientations to design VLEs and its effects on the learners is a key aspect of CBI instruction theory and plays a central role when the learning process is analyzed as a whole. Nevertheless, this stream of research is beyond the scope of this study, but is considered to be included in the next step of the present research project.

In consequence, this paper is going to be centred on a revision of the literature related to adoption and use of technology on learning contexts, comprising theories based on attitudes, motivation, and cognitive traits. The central objective is to propose a coherent synthesis of them, which will guide the research hypothesis related to the role of cognitive characteristics, VLE characteristics, and the interaction among them to induce attitudes towards learning technology and its adoption.

2 The adoption of Virtual Learning Environments.

Most of the investigation on adoption has been focused on two types of variables, in part depending on the disciplinary approach of the researchers. On one hand, a stream has been focused on the attitudes of learners (and teachers) towards technology and their reactions to the design of the VLE and learning objects. In the other hand, the focus has been set on individual cognitive characteristics that seem to be related to preferences and usage styles, in particular behavioural drivers and styles of behavioural regulation.

2.1 Attitudes and perceptions

The role of *attitudes and thoughts* on the adoption of computer technology has been deeply explored under the perspective of work and entrepreneurial organizations since the early 80´s. Under this stream of research, the use of new information technology was expected to be followed by an improvement on productivity, reason why the prediction of the use of a new technology was an expected and necessary information.

The study of variables related to behavioural intentions based on previous work from Social Cognitive theory [11], the paradigm of Cost-Benefit [12] , and the theory of Reasoned Behaviour [13], leads to the development of Technology Acceptance Model (TAM) [14] by Fred Davis. Davis' model purposes the idea that perceived usefulness and perceived ease of use directly influence the intention of use a determined technology. Strong evidence supports the core of TAM when is applied to

learning contexts [15]–[17], but several modifications have been introduced to adequate it to a variety of scenarios and conditions. One of the most recurrent is the inclusion of computer self-efficacy, a measure of how good the user think he/she is when using computer systems [18], [19], which can influence the adoption by reducing anxiety and stress, and creating a more pleasant experience.

Other important path of research has been centred on the instrumentality of the learning technology – the individual evaluation of how helpful it is – and the user's perceptions of it. It has been observed that when technology is perceived as adequate to fulfil a given task , people tend to prefer it instead of others and to use it more often [20], [21]. This theory – known as Task-Fit theory (TFT) – proposes that when technology is aligned with the task, the usage rate and the performance of the user improves. Even when the virtual environment or tool remains stable, if the task change, then the perception of fit may vary, hence it may suggests that the value of a tool is not stable, despite it is evaluated as a whole.

While Davis' model is based on affective aspects of adoption, the TFT relies on a more rational evaluation of the instrumentality of technology, therefore complementariness between them is far to be discarded. On the contrary, this research suggests that confluence of perspectives enriches the comprehension of the phenomena.

2.2 User learning characteristics

Another perspective to understand why people adopt virtual environments to learn is based on personal cognitive characteristics.

The individual characteristics that make learners different each other might include a wide diversity of aspects such as autonomy or information processing. While information processing and other related tasks are usually linked to performance, autonomy and motivational drivers appear as related to adoption and use of technology. Learner traits as academic locus of control and learning approach might be involved in the adoption process, making learners keen to some sorts of learning environments more than others do.

 Since the early work of Coovert & Goldstein linking locus of control with positives attitudes towards computers in work settings [22], it has been directly associated to technology and learning in different aspects. Intrinsic Locus of Control has been positively related to high motivation to use technology and with high performance [23]–[25], making learners more enthusiastic to be engaged with digital environments and its related activities. Nonetheless, a study conducted by Levy [26] on continuance of use of e-learning instruction found that locus of control is not related with dropout rates, suggesting that engagement with VLE is rather related with satisfaction and other attitudes. Therefore, it seems reasonable to think that locus of control might be involved in the initial stage of adoption (the behavioural intention) rather than in its continuance of use or engagement.

In addition to previous ideas, have been observed that student trends prior to the task might influence their intentions of use a determined VLE. Biggs [27] proposed a two–factors model of learning approach, identifying a surface and a deep learning profiles linked to extrinsic and intrinsic motivations respectively. This framework intents to be less context-dependent than other theories focused on information

processing and learning styles [28], [29]. Biggs' proposal shifts the focus from specific learning aspects to general trends and strategies that shape the behavior of the student through the learning process, mobilizing cognitive resources and energy. According to this, the surface approach is characterized by intentions to fulfil the given task with less effort and it is mobilized by a "fear of failure", while a deep approach is related to intrinsic interest, and a meaningful and effective learning experience.

A direct and significant relationship was found between the rate of use of VLE & performance to learning approach in a campus-based course complemented by learning technology [30]. Besides, satisfaction with a less directive learning design has been found to be related to a deep learning approach [31]. Nonetheless, the limited data on this matter forces us to be cautious and to test to what point these general drivers influence learners to adopt a determined VLE, and to investigate the relationship with other variables such as VLE design, autonomy, or attitudes towards technology.

3 Objectives and hypotheses

The purpose of the present research is to identify the variables related with the technology adoption process from an integrative e perspective that includes VLE characteristics, the learner characteristics, and the interaction between them. Specifically, the adoption process will be examined considering the characteristics of the learner and its evaluations on the learning environment. A second objective to be developed in future studies is to understand how these variables could be linked to effectiveness through the entire learning process.

Hypothesis I: Learning approach and Academic Locus of Control will influence learner attitudes towards the VLE and consequently, learner intention of use.

Hypothesis II: The design of the VLE will have a direct and significant effect on attitudes and intention of use.

Hypothesis III: Perceived Usefulness, Perceived Ease of Use, Perceived Fit and Computer Self-efficacy will be directly associated to Intention of use.

4 Methodology

4.1 Design and sample

This was a cross-sectional online study. Participants completed a measure of learning approach (2 dimensions, surface and deep approach) and academic locus of control. Then they were randomised to be allocated into either a structured or an unstructured virtual learning environment in which they should have to complete a hypothetical task. The simulated "structured VLE" included specific folders with reading resources and class materials, and everything was ordered as a traditional learning process logic. The "unstructured VLE" was the complete opposite, and it was only constituted by a

web browser, the more flexible tool to find the resources to complete the given task, allowing students to access to different kinds of materials depending of their interests and preferences.

Finally, the participants reported Perceived Usefulness (PU), Perceived Ease of Use (PEOU), Perceived Fit (PF), Computer Self-Efficacy (CSE) and Intention of Use (IOU) associated to the specific VLE they were allocated.

A total of 228 volunteers conformed the sample. All of them were undergraduate or postgraduate students of a single University.
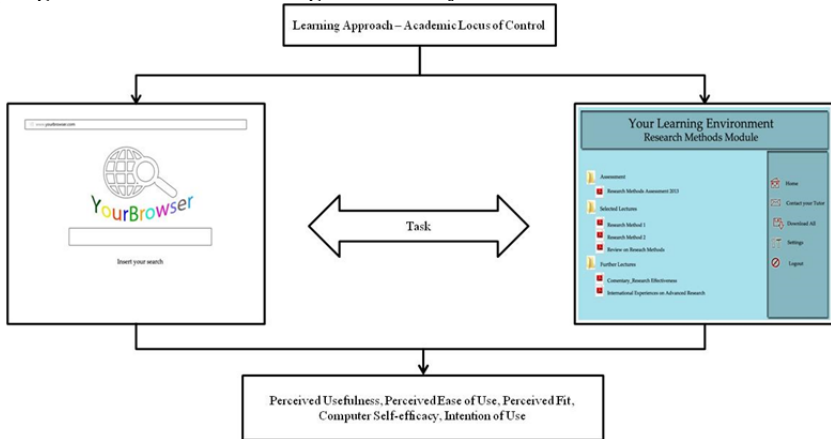
Figure 1. Research design, illustrating the initial measures, the randomizations of the participants, the two different VLE, and the second set of measures.

4.2 Instruments

A set of questionnaires and scales utilized in previous studies was selected to collect the required data. Then, the selected material was adapted to the conditions of the present study. Specifically, the scales of *perceived usefulness, computer self-efficacy* and *behavioural intention* were extracted from a study conducted for Liaw [38], *perceived ease of use* was extracted from a work by Sun [45], the inventory of *academic locus of control* from Levy [26], and the *learning approach* was assessed by the revised two-factor Study Process Questionnaire (R-SPQ-2F) developed by Biggs [27].

5 Results

5.1 Descriptives

The average age of the sample is 21.35 years old (S.D. =3.346). 67.1% of the 228 respondents were female (153), and 117 subjects (51.3%) evaluated the structured LE.

The effect of *gender* over all the research variables was assessed with a T-test to compare the means of each group, finding no significant differences between them.

Table 1 includes the mean, S.D. and internal consistency of each instrument used. Results are aligned with what has been found in previous studies.

It might be observed in table 2 that correlations shown no evidence of collinearity and that variables follow the rationale of the hypotheses, being possible to identify two blocks of variables: one composed by attitudes towards technology, and the other composed by learner traits.

The two-factors scale of learning approaches was confirmed by a factorial analysis.

Table 1. Mean, standard deviation, and internal consistency of the instruments used in the study.

	Items	Mean	S.D.	IC
1. Learning Environment				
2. Academic Locus of Control	12	37.86	5.931	0.743
3. Surface Approach	10	27.17	6.347	0.804
4. Deep Approach	10	32.16	5.872	0.786
5. Self-efficacy	3	12.31	2.259	0.894
6. Perceived Usefulness	3	12.34	1.777	0.899
7. Perceived Ease of Use	4	16.12	2.806	0.878
8. Perceived Fit	7	24.46	3.769	0.698
9. Intention of Use	3	11.93	1.811	0.739

Table 2. Partial correlations between the research variables.

	1	2	3	4	5	6	7	8	9
1. LE	1	-.021	.051	.049	.239**	.250**	.200**	.218**	.310**
2. ALC		1	-.326**	.358**	.057	.091	.151*	-.009	.049
3. SA			1	-.441**	-.068	-.065	-.071	.096	-.017
4. DA				1	.159*	.146*	.222**	.067	.128
5. CSE					1	.576**	.609**	.382**	.453**
6. PU						1	.517**	.589**	.570**
7. PEOU							1	.413**	.555**
8. PF								1	.562**
9. IOU									1

$*p<.05$ $**p<.001$
LE: Learning Environment **ALC**: Academic Locus of Control **SA**: Surface Approach **DA**: Deep Approach **CSE**: Computer Self-efficacy **PU**: Perceived Usefulness **PEOU**: Perceived Ease of Use **PF**: Perceived Fit **IOU**: Intention of Use.

5.2 Research Model

The research model was analyzed in AMOS 20 using Generalized Least Square method. To assess the quality of the model it was considered the X^2 index (p>.05 is good), the Goodness-of fit-index (GFI, a value higher than .900 is good), and the Comparative Fit Index (CFI, good if .95 or more).

The result of the analysis for the research model was X^2(9)=67.890, p=.000, GFI=.915, and CFI=.492, and several path could not be confirmed. Nonetheless, the direct effect of VLE on PU, PEOU, PF and IOU can be observed, the same for the direct effect of DA on PF, PEOU, and PU, and the indirect effect of VLE and on IOU can be observed as well (see table 3).

From previous, it can be sustained that all the three hypotheses received a partial support. Academic Locus of Control has no significant effect on the variables, and Learning Approach is relevant just in its Deep dimension. The design of the VLE appears as central in this research model, since it influences all the attitudes –as is logical to think –, but the design that enhance students positive attitudes is the unstructured VLE, in a significant way. Finally, the assess of the modified TAM shown that computer self-efficacy is not relevant – for this sample characteristics – as a predictor of intention of use, while the rest of the attitudes towards the VLE have a direct and significant effect on the intention of use.

Table 3. Standardized direct and indirect effect of the research variables.

Standardized Direct Effect		
PU <— VLE = .362	PU <— DA = .197	IOU <— PU = .169
PF <— VLE = .275	PF <— DA = .166	IOU <— PF = .248
PEOU <— VLE = .225	PEOU <— DA = .255	IOU <— PEOU = .295
IOU <— VLE = .159		
Standardized Indirect Effect		
IOU <— VLE = .196	IOU <— DA = .150	

6 Discussion

The conclusions of the present research come to present a more complex learner than was thought when facing a digital learning environment. On one hand, the findings of this study tell about a learner that is not the same facing different learning scenarios, that is going to mobilize its cognitive resources according to specific situations. Following this, the main challenges for learning planners and practitioners are two: One is to understand in a better way how the learning environment might influence the way learners drive their cognitive resources, how it motivates them, how it engages them into significant learning. The second is to think how to build such learning settings that on one hand get the best of those more curious, and on the other hand, assures a minimum of achievement for those that – for different reasons that escape to this study – are less enthusiastic into invest more effort that the adequate to fulfil a task. The main aspect to be considered is the strong effect of the VLE design on the outcomes. The attention must to keep being focused on the design of the VLE, its characteristics and especially how its design balances the constraint of an instructor/teacher-guided scheme with the liberty of the explicit lack of structure that gives liberty to the learner to self-guide its own learning process. It is important to notice that not all the students possess the same learning profile, the same goals, or the same strategies, so that the challenge is to engage them in a learning process that might could seen as relevant and coherent in accordance with their particular profiles.

On the other hand, data suggests that contrarily to what might be thought, academic locus of control and deep learning approach have no direct effect on the intention of use, and its role on this adoption model is minor. In fact, ALC have no significant effect on the variables, and DA is just indirectly involved in adoption. It might be inferred that the effect of a high DA is to produce a better disposition of the learners to evaluate in a positive way the learning conditions, maybe because their motivational driver is intrinsic and depends less of environmental conditions. However, it is important to take in account that university students compose the sample, therefore one might supposes they have a positive driver to learning activities.

Additionally, the importance of learner evaluations about the VLE is confirmed as the strongest and easily measurable predictor of intention of use. Nonetheless, the initial intention of use is not a warranty of continued use or of effective use. In education cannot be assumed that the mere use of a digital learning environment or learning software is going to improve the learning outcome just by being used. In fact, the relationship between adoption and performance may be influenced by other variables, and even the concept of *"effectiveness"* is not always conceptualize in an unequivocally way.

It was set before that there is not a common framework of CBI effectiveness and in most of the studies is very hard to observe a bridge between adoption and performance. The present model intends to set the basis to this by including variables such as ALC and learning approach on it. Even when their role in the adoption process seems to be weak, it may not be discarded an interaction between them and variables related to continuance of use and performance, such as engagement with the training course – in the case of learning approach –, or behavioural control – in the case of academic locus of control –, among others. Future research must take on

account this relationships, their effects and interactions, in order to find a way to improve the effectiveness of CBI (adoption and performance), which biggest challenge is not just being effective for those who already have experience with it, but with all the rest of new potential users.

Finally, the transition from intentions to actions, from the adoption of a VLE to the attainment of learning goals, is a key aspect of this research project. However, the difficulties start with the very definition of *effectiveness*, mainly because the success of a training course might include a large number of pedagogical, economic and social factors [32]. Facing a lack of common framework to understanding and measuring the effectiveness of VLE, including objective measures [33]–[36], learning self-perception [37], [38] , or a combination of various factors [19], [39]. This diversity makes hard to compare studies or obtain strong evidence on the effectiveness of CBI [40]. Up to date, not many studies have been conducted using this framework [42]–[44], but the results seem to be favourable, not just to the proposed relationship between the variables, but also to the explanatory power of the model. The next steps of this research is oriented to the integration of the framework on adoption of learning technology with the CBI theory in order to developing an integrated framework to improve effectiveness of VLEs.

References

1. M. K. O. Lee, C. M. K. Cheung, and Z. H. Chen, "Acceptance of Internet-based learning medium: the role of extrinsic and intrinsic motivation," *Inf. Manag.*, vol. 42, no. 8, pp. 1095–1104, 2005.
2. H.-J. Lai, "Information Literacy Training in Public Libraries: A Case from Canada," *Educ. Technol. Soc.*, vol. 14, no. 2, pp. 81–88, 2011.
3. R. Rivard, "Measuring the MOOC dropout rate," *Insid. High. Ed*, vol. 8, 2013.
4. D. Yang, T. Sinha, D. Adamson, and C. P. Rosé, "Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses," in *Proceedings of the 2013 NIPS Data-Driven Education Workshop*, 2013.
5. P. Lameras, P. Levy, I. Paraskakis, and S. Webber, "Blended university teaching using virtual learning environments: conceptions and approaches," *Instr. Sci.*, vol. 40, no. 1, pp. 141–157, Jan. 2012.
6. K. Samuelowicz and J. D. Bain, "Revisiting academics' beliefs about teaching and learning," *High. Educ.*, vol. 41, no. 3, pp. 299–325, Apr. 2001.
7. A. Roberts, A. D. Madden, and S. Corrall, "Putting Research into Practice: An Exploration of Sheffield iSchool Approaches to Connecting Research with Practice," *Libr. Trends*, vol. 61, no. 3, pp. 479–512, 2013.
8. P. Levy, O. Aiyegbayo, and S. Little, "Designing for inquiry-based learning with the Learning Activity Management System," *J. Comput. Assist. Learn.*, vol. 25, no. 3, pp. 238–251, Jun. 2009.
9. M. Kordaki, "High school computing teachers' beliefs and practices: A case study," *Comput. Educ.*, vol. 68, no. 0, pp. 141–152, Oct. 2013.
10. P. Levy, "A methodological framework for practice-based research in networked learning," *Instr. Sci.*, vol. 31, no. 1–2, pp. 87–109, Jan. 2003.
11. A. Bandura, *Social foundations of thought and action : a social cognitive theory.* London: Prentice-Hall, 1986.
12. L. R. Beach and T. R. Mitchell, "A Contingency Model for the Selection of Decision Strategies," *Acad. Manag. Rev.*, vol. 3, no. 3, pp. 439–449, 1978.

13. I. Ajzen, *From intentions to actions: A theory of planned behavior*. Springer, 1985.

14. F. D. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *Mis Q.*, vol. 13, no. 3, pp. 319–340, 1989.

15. J. Schepers and M. Wetzels, "A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects," *Inf. Manag.*, vol. 44, no. 1, pp. 90–103, 2007.

16. B. Sumak, M. Hericko, and M. Pusnik, "A meta-analysis of e-learning technology acceptance: The role of user types and e-learning technology types," *Comput. Human Behav.*, vol. 27, no. 6, pp. 2067–2077, 2011.

17. M. Turner, B. Kitchenham, P. Brereton, S. Charters, and D. Budgen, "Does the technology acceptance model predict actual use? A systematic literature review," *Inf. Softw. Technol.*, vol. 52, no. 5, pp. 463–479, 2010.

18. I. H. Hwang, C. C. Chang, H. L. Wang, S. J. Tsai, and T. Y. Chen, "Mediating Effects of Computer Self-efficacy between Learning Motivation and Learning Achievement," in *Education Management, Education Theory and Education Application*, vol. 109, Y. Z. Wang, Ed. Berlin: Springer-Verlag Berlin, 2011, pp. 67–73.

19. R. D. Johnson, S. Hornik, and E. Salas, "An empirical examination of factors contributing to the creation of successful e-learning environments," *Int. J. Hum. Comput. Stud.*, vol. 66, no. 5, pp. 356–369, 2008.

20. D. L. Goodhue and R. L. Thompson, "Task-technology fit and individual performance," *MIS Q. Manag. Inf. Syst.*, vol. 19, no. 2, pp. 213–233, 1995.

21. W. S. Lin, "Perceived fit and satisfaction on web learning performance: IS continuance intention and task-technology fit perspectives," *Int. J. Hum. Comput. Stud.*, vol. 70, no. 7, pp. 498–507, 2012.

22. M. D. Coovert and M. Goldstein, "LOCUS OF CONTROL AS A PREDICTOR OF USERS ATTITUDE TOWARD COMPUTERS," *Psychol. Rep.*, vol. 47, no. 3, pp. 1167–1173, 1980.

23. W. Eom and R. A. Reiser, "The effects of self-regulation and instructional control on performance and motivation in computer-based instruction," *Int. J. Instr. Media*, vol. 27, no. 3, p. 247, 2000.

24. Y. J. Joo, S. Joung, and W. J. Sim, "Structural relationships among internal locus of control, institutional support, flow, and learner persistence in cyber universities," *Comput. Human Behav.*, vol. 27, no. 2, pp. 714–722, 2011.

25. A. Broos and K. Roe, "The digital divide in the playstation generation: Self-efficacy, locus of control and ICT adoption among adolescents," *Poetics*, vol. 34, no. 4–5, pp. 306–317, 2006.

26. Y. Levy, "Comparing dropouts and persistence in e-learning courses," *Comput. Educ.*, vol. 48, no. 2, pp. 185–204, 2007.

27. J. Biggs, D. Kember, and D. Y. P. Leung, "The revised two-factor Study Process Questionnaire: R-SPQ-2F," *Br. J. Educ. Psychol.*, vol. 71, pp. 133–149, 2001.

28. F. Coffield, D. Moseley, E. Hall, and K. Ecclestone, "Learning styles and pedagogy in post-16 learning: A systematic and critical review," 2004.

29. F. Romanelli, E. Bird, and M. Ryan, "Learning styles: A review of theory, application, and best practices," *Am. J. Pharm. Educ.*, vol. 73, no. 1, 2009.

30. R. Ellis, M. Weyers, and J. Hughes, "Campus-based student experiences of learning technologies in a first-year science course," *Br. J. Educ. Technol.*, vol. 44, no. 5, pp. 745–757, 2013.

31. E. Gurpinar, E. Kulac, C. Tetik, I. Akdogan, and S. Mamakli, "Do learning approaches of medical students affect their satisfaction with problem-based learning?," *Adv. Physiol. Educ.*, vol. 37, no. 1, pp. 85–88, 2013.

32. P. M. Halachev, "E-learning effectiveness," *Interactive Computer Aided Learning (ICL) - International Conference*. Austria, pp. 551–557, 2009.

33. S. Kekkonen-Moneta and G. Moneta, "E-learning in Hong Kong: comparing learning outcomes in online multimedia learning and lecture versions of an introductory computing course," *Br. J. Educ. Technol.*, vol. 33, no. 4, pp. 423–433.

34. H. Lim, S.-G. Lee, and K. Nam, "Validating E-learning factors affecting training effectiveness," *Int. J. Inf. Manage.*, vol. 27, no. 1, pp. 22–35, 2007.

35. M. Merkt, S. Weigand, A. Heier, and S. Schwan, "Learning with videos vs. learning with print: The role of interactive features," *Learn. Instr.*, vol. 21, pp. 687–704.

36. P. Stonebraker and J. Hazeltine, "Virtual learning effectiveness," *Learn. Organ.*, vol. 11, no. 3, pp. 209–225.

37. N. Buzzetto-More and B. Mitchell, "Student performance and perceptions in a web-based competitive computer simulation," *Interdiscip. J. E-Learning Learn. Objects*, vol. 5, no. 1, pp. 73–90, 2009.

38. S.-S. Liaw, "Investigating students' perceived satisfaction, behavioral intention, and effectiveness of e-learning: A case study of the Blackboard system," *Comput. Educ.*, vol. 51, no. 2, pp. 864–873, 2008.

39. W. Bhuasiri, O. Xaymoungkhoun, H. Zo, J. J. Rho, and A. P. Ciganek, "Critical success factors for e-learning in developing countries: A comparative analysis between ICT experts and faculty," *Comput. Educ.*, vol. 58, no. 2, pp. 843–855, 2012.

40. T. Sitzmann, K. Kraiger, D. Stewart, and R. Wisher, "The comparative effectiveness of web-based and classroom instruction: A meta-analysis," *Pers. Psychol.*, vol. 59, no. 3, pp. 623–664, 2006.

41. J. S. Lowe and E. F. Holton, "A theory of effective computer-based instruction for adults," *Hum. Resour. Dev. Rev.*, vol. 4, no. 2, pp. 159–188, 2005.

42. E. M. Ineson, T. Jung, C. Hains, and M. Kim, "The influence of prior subject knowledge, prior ability and work experience on self-efficacy," *Leis. Sport Tour. Educ.*, vol. 12, pp. 59–69, 2013.

43. H. J. Lai, "The Influence of Adult Learners' Self-Directed Learning Readiness and Network Literacy on Online Learning Effectiveness: A Study of Civil Servants in Taiwan," *Educ. Technol. Soc.*, vol. 14, no. 2, pp. 98–106, 2011.

44. O. N. Shinkareva and A. Benson, "The Relationship between Adult Students' Instructional Technology Competency and Self-Directed Learning Ability in an Online Course," *Hum. Resour. Dev. Int.*, vol. 10, no. 4, pp. 417–435, 2007.

45. P.-C. Sun, R. J. Tsai, G. Finger, Y.-Y. Chen, and D. Yeh, "What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction," *Comput. Educ.*, vol. 50, no. 4, pp. 1183–1202, 2008.

Towards a Topic Detection and Tracking Framework for Albanian News

Klesti Hoxha

University of Tirana, Faculty of Natural Sciences, Department of Computer Science,
Blv. Zogu I, Nr. 25/1, 1001 Tirana, Albania
klesti.hoxha@fshn.edu.al

Abstract. The number of online news sources that publish news articles written in Albanian has been increased recently. This has led to an increase of information overload due to news article mirroring and quickly produced news event updates published by many sources. In this paper it is presented a preliminary proposal of a highly independent framework for topic detection and tracking that deals with news written in Albanian. It is supposed to be used as a separate module to existing news search engines, or to serve as a backend for various desktop, mobile and subscription based notification interfaces. An initial implementation idea that makes use of existing open source packages is also presented.

Keywords: topic tracking, information retrieval, news search engine, recommender system, personalized information

1   Introduction

Most of the Albanian news producing entities (agencies, newspapers, and radio/television broadcasters) publish their news online [1]. This has led to an increase of readers that use these online sources as a first choice. They are situated in Albania, the Albanian speaking communities in its neighboring countries and in many other countries over the world (the Albanian Diaspora). Even though the number of news produced in Albania is not too big in comparison with various global news sources, there is still a big information overload. Many events are covered by different media and sometimes the exact same news is published in different news providers (maybe originating by the same source, i.e. a news agency). News search engines can facilitate this process, but international ones that do not support the Albanian language cannot help the news retrieval process. Also various users are interested in following a live coverage of a particular event reported in the news, or real-time updates of news about a specific topic. The removal of duplicated information and a concentrated topic-based way of news browsing would further improve the user experience of a software based news recommender [29].

While most of the commercial news search engines and the prototypes proposed in academia [14] allow the automatic categorization of news among some predefined

topics (i.e. technology, sports, politics), this is not enough for reducing the information overload. Readers may be interested about a particular topic emerged recently (i.e. a fire in the forests of Albania) and get quick updates about it. Some news topics may evolve over a relatively long period of time, for example news coverage about Malaysia Airlines flight 370's disappearance[1] lasted for about one month. Dealing with emerging topics creates the need for automatic ways for topic detection and tracking.

Topic detection and tracking regarding news events has been covered by many researchers worldwide [7, 15, 28, 35]. Yang et al. [35] point out the need of an intelligent system that automatically detects new events (uncovered news) and marks them as new topics. It should also be possible to allow the tracking of updates regarding a previously identified topic. The basic approach of topic detection is by clustering the news data stream into groups of related news [9]. However the dynamic nature of news publishing includes several news updates or new reports about the same event. These updates may be done over a long period. Therefore the news clustering needs to allow for an incremental addition of items, adding newly published news in previous detected clusters [4].

While an extensive amount of topic detection and tracking research about news articles written in English or other languages has been documented, to the best of our knowledge there has been no reported work about this area that deals with news articles written in Albanian. In this paper it is presented a first proposal of a topic detection and tracking framework for news articles written in Albanian. The initial proposed implementation relies on well known open source packages that deal with information retrieval and text mining tasks. It also covers a news collection and indexing strategy that enables the functioning of many topic detection (mining) techniques [9, 21]. The proposed framework will serve as a pillar for the design and implementation of a news search engine that delivers personalized [20, 22] news articles written in Albanian to interested readers.

2   Design Goals

The main aim of this work is to define the architecture and initial implementation proposals of a topic detection and tracking framework that deals with news articles written in Albanian. The framework has been designed by considering the following design goals:

- It should be extensible and configurable by allowing the usage of different topic tracking and detecting algorithms.
- It should make possible the inclusion of natural language processing tools and features, like lexical databases [2] or named entity recognition [32].
- It should index news articles gathered through various data sources. This index should be offered through a separate layer (module) independent from the topic tracking part.

---

[1] http://en.wikipedia.org/wiki/Malaysia_Airlines_Flight_370

- It should allow for user based personalization through external user model inputs. Examples include news source or topic selection, similarity score threshold, usage data, etc.
- It should offer an experimentation environment about topic detection and tracking of Albanian news

The framework proposed below tries to meet all the described design goals.

3   Related Work

News retrieval systems have been proposed both as commercial products[2] and research prototypes. In this section is presented a brief review of the current state of research regarding news retrieval and recommendation.

Gulli describes a general anatomy of a news search engine [14]. He proposes a retriever module that collects news through syndication mediums like RSS and ATOM. The news gathered by the retriever are indexed and ranked through separate modules. Gulli proposes a simple tf-idf indexing that indexes news based on the frequencies of particular terms (words) in them. The term frequency index is used by the classifier and clustering engine. The classifier engine classifies news in different already defined categories, like sports, technology, etc. In this work a naïve Bayesian classifier is used that works in training mode for already manually classified news (in the actual feeds) or classifying mode if the news have not been manually categorized. The clustering engine tries to group similar news together in order to decrease the visualization of mirrored news from various sources. The algorithm used for clustering is k-means with a distance threshold. The search engine proposed by Gulli includes also a personalization services that sends by email recent news about a search query defined by the system's users. In an associated paper, Corso, Gulli and Romani present the ranking algorithm used by the above described system [10]. It uses topic based clustering and it takes into consideration also the time of publication and the importance of the news source. The proposed algorithm ranks not only the news articles, but also the sources of the news based on some given heuristics.

In [21] is described another RSS based news aggregator that clusters similar news together and represents them to the users. It uses an incremental story clustering that involves a flexible time window. There are no predefined categories. Topics "emerge" by the usage of a clustering algorithm that is again k-means. The proposed service is subscription-based; users have to subscribe to a group of RSS-feeds. A similar approach is described in [4]. The authors incrementally cluster news gathered from user-provided RSS feeds. Again a customized version of k-means clustering algorithm has been used and evaluated.

Katakis and Tsoumakas [19] also describe a RSS based news dissemination system. It offers personalization features regarding the user interests. There is an individual personalization for each feed that the user is subscribed to. The proposed

---

[2] Examples of news retrieval commercial products:
   Classic news search engines: https://news.google.com/ , http://news.search.yahoo.com/
   Social news recommender: https://www.linkedin.com/company/pulse-news

system includes an aggregator module that uses machine-learning algorithms to classify news offered by each subscribed feed as interesting or not interesting to a particular user. It also supports topic-based personalization; it allows users to monitor news regarding a specific topic that they are interested in. Similar to the system described by Gulli in [14], it also includes a notification module that notifies users by email about updates that might be interesting to them. The authors have also described the implementation details of this system. It has been built using open source technologies like PHP and MySQL.

Paliouras et al. describe another personalized news aggregator [26]. Differently from the systems described above, it collects news also from HTML sources, not only RSS or ATOM feeds. It offers a high level of personalization done through a separate personalization server that manages user model creation and helps the recommendation process. In this work the authors have compared different techniques for personalized news delivery that are based on the two basic types of recommender systems: content-based and collaborative filtering [3]. This system did not include similar items clustering. Their results did not show noticeable improvement of recommendation when using collaborative filtering techniques (when the recommendation is based on preferences of users similar to a particular user).

Other than offering personalized news and reducing information overload by clustering similar news items, attempts have been made to automatically summarize a cluster of related news items by describing the most important information of them. Such an approach is used in NewsInEssence [29]. It also includes a topic detection and tracking module that groups the downloaded articles by topic. Topics can also be based on user created clusters: users can create them by providing a seed article. Kastner and Monz describe another approach for key facts extraction from a particular news article [18]. It makes use of a combination of statistical and linguistic features. Two other news summarization systems have been described in [8] and [33].

Regarding the similarity measure used for comparing news in recommendation (related articles) and incremental clustering contexts, other than the traditional tf-idf weighting scheme based on the frequencies of the terms, have also been proposed schemes that make use of the similarity between concepts. Such an approach is presented by Bergamaschi et al. in [5]. It uses the similarity between concepts for calculating the similarity between the titles of two news articles. The heuristics for calculating the similarity between concepts makes use of WordNet [2], a well known lexical database of English. A similar approach is proposed also in [34]. Even though the results achieved by these two works are promising, they are not feasible for languages that do not have a large electronic dictionary or thesaurus that can be used for the calculation of similarity between concepts.

The method of choice for similarity calculation in topic detection and tracking is a customized traditional tf-idf weighting scheme [4, 7, 9, 28]. It uses the bag-of-words model for document representation and the cosine similarity [23] for calculating the actual similarity. Bogers and van den Bosch [6] further compare different similarity schemes for news recommendation. They conclude that for shorter texts the tf-idf scheme performs pretty well, however for long texts (i.e. when the whole article has been considered for similarity calculation) probability based retrieval schemes seem to work better.

Considering the dynamics of news data streams and the fact that updates to a previously reported news or other interpretations to it may be published with the passing of time, some works have been published that propose models that make use of this setting. NewsJunkie [13] is a system that provides personalized news feeds to users by pointing out the novelty of the newly published news. In another work, Shahaf and Guestrin [31] have proposed a method for automatically detecting information chains that show the coherence of a particular news story through time. Demartini et al. propose a system for automatic entity (people, locations, products, etc.) summarization based on a given news and other related ones published before it [11]. As defined by the authors, entity summarization refers to the extraction of a set of entities that better describe a document (i.e. a news article). The temporal distribution of documents is used by Peetz et al. [27] for relevant terms extraction. These terms are used for improving the queries used in a probabilistic language model retrieval approach. The latest have been shown to produce good results in news retrieval and recommendation contexts [6].

To the best of our knowledge there have been no previously reported works related specifically to topic detection and tracking about news articles written in Albanian. Therefore particular implications that might occur in this case are not clear. In the following paragraphs are described two concrete works that deal with clustering and recommendation of documents written in Albanian.

In [17] Kadriu presents a system for Albanian news clustering. It clusters crawled news into predefined categories, an approach similar to the one described by Gulli [14]. The author has not used any text preprocessing (stemming, stop word removal) specific to the Albanian language. The indexing scheme is again tf-idf. The hierarchical clustering algorithm that was used in this work produced good results (in terms of precision and recall) for a given set of categories (top news, world, sport, economy, technology, culture, showbiz).

A recommender system that deals with scientific articles written in Albanian is proposed in [16]. Even though the context is different and scientific articles have a predefined structure that is not always found in news articles, the results show a successful usage of a tf-idf weighting scheme for cosine measure based similarity calculation of documents written in Albanian.

Regarding commercial products, recently there has been an emergence of news search engines that deal with articles written in Albanian.[3] They offer a predefined news categorization (i.e. sports, politics) and no personalization or topic tracking features. Regarding news clustering, they group similar articles into related articles, reducing so the mirroring and overload of information.

4   Proposed framework

In this section it is presented a preliminary proposal of the topic detection and tracking framework together with some initial implementation ideas. The design goals

---

[3] Commercial search engines about news articles written in Albania:
 http://www.time.al
 http://www.grid.al

described in Section 2 of this article have been used for decision making regarding the modularity of the architecture and its concrete implementation.

Basically the proposed framework would make possible the creation of a news search engine with news recommendation features. It will index news crawled from various sources allowing for easy retrieval of them. News can be searched by a user specified query or recommended based on user profile data.

The main aim of this work in progress is to provide an experimental framework with loosely coupled modules in order to find out the settings that provide better results when dealing with retrieval tasks about news written in Albanian. The current available solutions are commercial ones and their internals have not been reported in academic papers. Many similar solutions make use of natural language processing tools for improving the indexing and retrieval of news. Because the support of such tools for the Albanian language is still limited, the initial plan is to use e traditional tf-idf ranking scheme possible together with a stemming algorithm and stop words removal. However, the architecture of the proposed framework allows the easy integration of additional NLP tools if they became available.

Another possible outcome of this work is the actual contribution to the improvement of natural language processing support for Albanian. The gathered news can create a corpus that can help on the annotation of words using supervised and unsupervised techniques. Furthermore the created dataset can serve as an experimental basis for further work on information retrieval systems that deal with documents written in Albanian.

4.1   System architecture

In Figure 1 is shown the system context diagram of the proposed framework. The framework is proposed to be independent from the news crawlers, user profiles and user management. It should allow external users to subscribe for personalized news updates delivered to them in different ways (email, mobile push updates, etc.). The personalization requests can vary from source choices to usage tracking for collaborative filtering recommendation [22, 24, 26]. The framework itself offers interfacing capabilities to external systems and can also be used by third party news search engines that need topic tracking and detection capabilities.
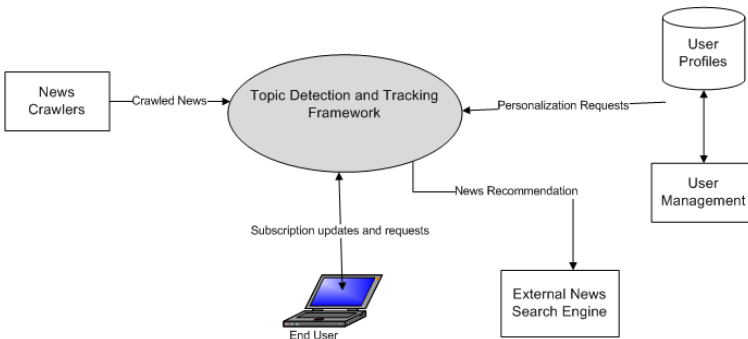
Fig. 1. System Context Diagram

In Figure 2 it is shown the internal architecture of the proposed framework. The modular structure facilitates the extensibility and configuration. It is possible to experiment with various algorithms for topic detection and tracking tasks. In the following paragraphs is provided a brief description of each proposed module.

The intermediate representation module offers an intermediate representation of the crawled news data. This can be stored in various formats, i.e. XML or to a relational database. The news crawled from an external crawler is processed by this module and stored for later indexing. This intermediate representation layer allows the usage of various indexing strategies and tools. It is also responsible for duplicate news detection. It should not allow the indexing of a news article (offered by a particular source) crawled and indexed in the past. If a particular news has been updated this change may need to be reflected in the actual index. Together with the extracted data it is stored the URL of the article and any additional metadata (crawled date, publication date, authors, manually specified categories, etc.).

The proposed framework makes use of *two indexes*: a classic tf-idf based index and a topic index that stores information about the generated topic clusters and allows for traceability.

The *tf-idf index* will be used for similarity calculation and incremental topic clustering. Even though recently there has been a publication of an Albanian version of WordNet (AlbaNet)[4], it is still very limited in comparison to the original English one. For this reason we plan to limit the indexing at this stage to a pure tf-idf scheme and not include concept based similarity calculation [5, 34]. Nevertheless the classic term frequency based clustering approach has been proved successful for many previous works (see Section 3). When indexing the term frequencies, a word stemmer [30] can be used in order to improve similarity calculation results.

The *topic index* stores the topics and their respective related news (clusters). It can be implemented in various ways, but a simple implementation would consist of a relational database.

The *topic detection* module is responsible for creating the topic index. It uses the intermediate representation of the news, the tf-idf index and the natural language processer module for topic detection and tracking. It consists of a clustering engine and a new event detection tool. Various algorithms can be implemented in order to achieve these latest functionalities. Regarding the similarity measure it consists of a scheme based on tf-idf indexing, for example the cosine similarity [23]. It can offer personalized topic detection if data about the user profiles are provided (externally). Other than classifying news in predefined categories like in [14], we plan to detect also emerging topic by using an incremental clustering scheme that uses a first story as the seed of a topic [4, 29]. The most representing words for a topic can be identified by using a similar approach to the ones reported in [18] and [29].

---

[4] http://fjalnet.com/
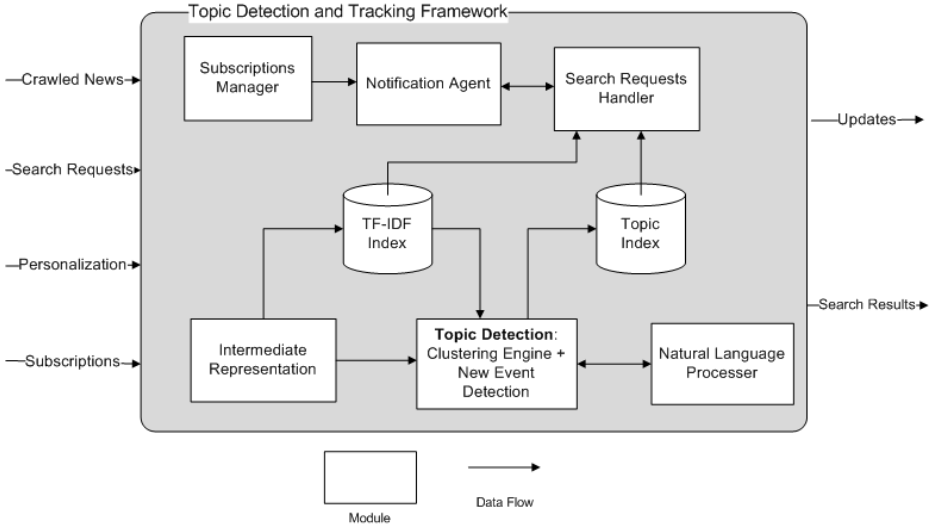
Fig. 2. The internal architecture of the proposed framework.

The *natural language processer module* can be used for enabling topic detection algorithms that rely on named entity recognition [32] or part of speech tagging. As NLP facilitators AlbaNet[5] and BabelNet [25] can be used. Their support of the Albanian language is still limited, however some initial insight can still be gained.

The *subscriptions manager* and *notification agent* modules manage topic tracking subscriptions by external users and are responsible for the actual notification sending. Notifications can be offered in several ways, including email, SMS, social updates, or mobile push notifications.

The *search requests handler* handles the external search requests or the ones originating from the notification agent. It produces search results by using the tf-idf index and/or the topic index based on the actual query.

4.2 Initial implementation ideas

In this section is proposed an initial implementation scheme that makes use of open source software packages. The modular architecture of this particular framework makes possible the usage of a mixture of technologies that communicate through wrappers or standard protocols.

The tf-idf index is suggested to be implemented with the Apache Lucene[6] framework. It has been used successfully for indexing and clustering news articles in some other projects [8, 12, 15, 33]. It offers keyword based search and its term-document vectors can be used for our concrete clustering and topic detection needs. It is also possible to override the default implementation of Lucene indexing, by specifying custom word stemmers. This holds also for the ranking of search results. The default usage of Lucene is done through Java, but wrappers are available, like for

---

[5] http://fjalnet.com

[6] http://lucene.apache.org/core/

example a Python wrapper[7]. The availability of Python wrappers for Lucene makes possible the usage of other tools written in Python. Another option would be Apache Solr[8], an extended search server based on Lucene. It is also supported by various wrappers including Python ones.

Apache Mahout[9] can be used for aiding the topic detection and tracking tasks. It offers implemented algorithms about topic mining, document clustering and classification. A possible machine learning library that is available for Python is scikit-learn[10].

The intermediate representation of news together with the topic index can be implemented by using a relational database, for example MySQL[11], or as a structured document like XML. The natural language processor can be implemented by using an open source package written in Java or Python, depending on the language of implementation of the topic tracker. Good candidates are Apache OpenNLP[12] and Python NTLK[13].

Many of the online Albanian news publishers do not include RSS or Atom Feeds [1]. In order to not exclude any sources, it is needed a custom HTML parsing for each provider that does not have a news feed. Scrapy[14] is a Python framework for web scrapping and crawling. It allows the extraction of information from XML (including RSS/Atom) and HTML. In combination with Scrapyd[15] it can be used for automated news retrieval (crawling). It offers many mechanisms that facilitate data scrapping and transformation therefore can be a good alternative for retrieving the news and providing them to the intermediate representation module described above.

5   Discussion

The architecture of the proposed framework is similar to previously reported ones [9, 26, 14, 21]. However the main intention of this work is to provide a highly configurable topic detection and tracking framework that can be used by news search engines or news recommenders. It offers a loosely coupled architecture that allows the integration of different modules not necessarily implemented with the same technology.

There are no previously reported works that deal with news articles written in Albanian, therefore we plan to produce useful insight regarding this research question. It is not completely known weather existing information retrieval techniques deal with a considerable performance with documents written in Albanian.

---

[7] http://lucene.apache.org/pylucene/

[8] http://lucene.apache.org/solr/

[9] https://mahout.apache.org/

[10] http://scikit-learn.org/

[11] http://www.mysql.com

[12] http://opennlp.apache.org/

[13] http://ntlk.org

[14] http://scrapy.org/

[15] http://scrapyd.readthedocs.org/en/latest/

The principal contribution of this work will be a highly configurable intelligent news search engine that will facilitate the quick and easy retrieval of news articles written in Albanian. This will be achieved by means of the proposed topic detection and tracking framework. Topic detection and tracking also serves as facilitator in other contexts (e-commerce, e-health systems, etc.), therefore the results of this work can create a good insight of the feasibility and best practices that needs to be followed when dealing with topic detection and tracking of Albanian documents. The created news dataset can also be used for various future works about natural language processing of Albanian.

When the system is ready it will be evaluated with two different methodologies. The primary evaluation will be user driven. It will use the feedback of users for calculating precision and recall [23]. The second planned evaluation is a comparison with existing English datasets or topics detected in automatically translated documents.

6   Conclusions

While the increase of Albanian online news providers has created better opportunities for live coverage of news events, the information overload due to mirrored news and frequent updates makes it a difficult process. Topic detection and tracking has been proposed for facilitating the task of new event (freshly reported news) detection and topic evolution over time. While extensive work has been done for news articles written in English and some other language, to the best of our knowledge there are no reported works that deal with news articles written in Albanian.

Incremental news clustering [4] is the method of choice for many proposed topic detection algorithms [9]. It dynamically enlarges previously detected clusters (topics) with new articles published with the passing of time. If no previous cluster has been created, a news article can serve as a topic seeder. While several similarity measures have been proposed, a classic tf-idf based indexing scheme with a cosine similarity based similarity measure has been proven to be successful as a clustering enabler [34].

In this paper it is proposed a preliminary architecture for a topic detection and tracking framework that deals with news articles written in Albanian. It is designed as a separate module that can be incorporated in existing news search engines or used as a backend for a subscription based topic alert service (a notification is sent to subscribers when a new article about a particular topic has been published). Together with the preliminary architecture some implementation ideas that make use of existing open source software packages have been presented.

References

1. List of Albanian news publishing websites, http://shqiperia.com/faqe/lajme-media
2. Princeton University: About WordNet (2010), http://wordnet.princeton.edu

3. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. Knowledge and Data Engineering, IEEE Transactions on 17(6), 734–749 (2005), http://ieeexplore.ieee.org/xpls/-abs_all.jsp?arnumber=1423975

4. Azzopardi, J., Staff, C.: Incremental clustering of news reports. Algorithms 5(3), 364–378 (2012)

5. Bergamaschi, S., Guerra, F., Orsini, M., Sartori, C., Vincini, M.: Relevant news: a semantic news feed aggregator. SEMANTIC WEB APPLICATIONS AND PERSPECTIVES (SWAP 2007) p. 150 (2007)

6. Bogers, T., Van den Bosch, A.: Comparing and evaluating information retrieval algorithms for news recommendation. In: Proceedings of the 2007 ACM conference on Recommender systems. pp. 141–144. ACM (2007), http://dl.acm.org/citation.cfm?id=1297256

7. Can, F., Kocberber, S., Baglioglu, O., Kardas, S., Ocalan, H.C., Uyar, E.: New event detection and topic tracking in turkish. Journal of the American Society for Information Science and Technology 61(4), 802–819 (2010), http://onlinelibrary.wiley.com/doi/-10.1002/asi.21264/full

8. Chen, C.M.: Intelligent location-based mobile news service system with automatic news summarization. Expert Systems with Applications 37(9), 6651–6662 (2010), http://-www.sciencedirect.com/science/article/pii/S0957417410002368

9. Chung, S., McLeod, D.: Dynamic topic mining from news stream data. In: On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, pp. 653–670. Springer (2003), http://link.springer.com/chapter/10.1007/978-3-540-39964-3_42

10. Del Corso, G.M., Gull, A., Romani, F.: Ranking a stream of news. In: Proceedings of the 14th international conference on World Wide Web. pp. 97–106. ACM (2005), http://-dl.acm.org/citation.cfm?id=1060764

11. Demartini, G., Missen, M.M.S., Blanco, R., Zaragoza, H.: Entity summarization of news articles. In: Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval. pp. 795–796. ACM (2010), http://dl.acm.org/-citation.cfm?id=1835449.1835620

12. Fernandez, N., Blazquez, J.M., Fisteus, J.A., Sanchez, L., Sintek, M., Bernardi, A., Fuentes, M., Marrara, A., Ben-Asher, Z.: News: Bringing semantic web technologies into news agencies. In: The Semantic Web-ISWC 2006, pp. 778–791. Springer (2006), http://-link.springer.com/chapter/10.1007/11926078_56

13. Gabrilovich, E., Dumais, S., Horvitz, E.: Newsjunkie: providing personalized newsfeeds via analysis of information novelty. In: Proceedings of the 13th international conference on World Wide Web. pp. 482–490. ACM (2004), http://dl.acm.org/citation.cfm?id=988738

14. Gulli, A.: The anatomy of a news search engine. In: Special interest tracks and posters of the 14th international conference on World Wide Web. pp. 880–881. ACM (2005), http://-dl.acm.org/citation.cfm?id=1062778

15. Ha-Thuc, V., Mejova, Y., Harris, C., Srinivasan, P.: A relevance-based topic model for news event tracking. In: Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval. pp. 764–765. ACM (2009), http://-dl.acm.org/citation.cfm?id=1572117

16. Hoxha, K., Kika, A., Gani, E., Greca, S.: Towards a modular recommender system for research papers written in Albanian. International Journal of Advanced Computer Science and Applications(IJACSA) 5(4) (2014)

17. Kadriu, A.: Albanian news clustering. In: Proceedings of the 4th International Conference on Information Systems and Technology Innovations (2013)

18. Kastner, I., Monz, C.: Automatic single-document key fact extraction from newswire articles. In: Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics. pp. 415–423. Association for Computational Linguistics (2009), http://dl.acm.org/citation.cfm?id=1609113

19. Katakis, I., Tsoumakas, G., Banos, E., Bassiliades, N., Vlahavas, I.: An adaptive personalized news dissemination system. Journal of Intelligent Information Systems 32(2), 191–212 (2009), http://link.springer.com/article/10.1007/s10844-008-0053-8

20. Li, L., Zheng, L., Yang, F., Li, T.: Modeling and broadening temporal user interest in personalized news recommendation. Expert Systems with Applications 41(7), 3168–3177 (2014), http://www.sciencedirect.com/science/article/pii/S0957417413009329

21. Li, X., Yan, J., Deng, Z., Ji, L., Fan, W., Zhang, B., Chen, Z.: A novel clustering-based rss aggregator. In: Proceedings of the 16th international conference on World Wide Web. pp. 1309–1310. ACM (2007), http://dl.acm.org/citation.cfm?id=1242824

22. Liu, J., Dolan, P., Pedersen, E.R.: Personalized news recommendation based on click behavior. In: Proceedings of the 15th international conference on Intelligent user interfaces. pp. 31–40. ACM (2010), http://dl.acm.org/citation.cfm?id=1719976

23. Manning, C.D., Raghavan, P., Schütze, H.: Introduction to information retrieval, vol. 1. Cambridge university press Cambridge (2008)

24. Montes-Garcá, A., Alvarez-Rodrguez, J.M., Labra-Gayo, J.E., Martnez-Merino, M.: Towards a journalist-based news recommendation system: The wesomender approach. Expert Systems with Applications 40(17), 6735–6741 (2013), http://www.sciencedirect.com/science/article/pii/S0957417413004272

25. Navigli, R., Ponzetto, S.P.: Babelnet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. Artificial Intelligence 193, 217–250 (2012), http://www.sciencedirect.com/science/article/pii/S0004370212000793

26. Paliouras, G., Mouzakidis, A., Moustakas, V., Skourlas, C.: Pns: A personalized news aggregator on the web. In: Intelligent interactive systems in knowledge-based environments, pp. 175–197. Springer (2008), http://link.springer.com/chapter/10.1007/978-3-540-77471-6_10

27. Peetz, M.H., Meij, E., de Rijke, M., Weerkamp, W.: Adaptive temporal query modeling. In: Advances in Information Retrieval, pp. 455–458. Springer (2012), http://link.springer.com/chapter/10.1007/978-3-642-28997-2_40

28. Pon, R.K., Cardenas, A.F., Buttler, D., Critchlow, T.: Tracking multiple topics for finding interesting articles. In: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 560–569. ACM (2007), http://dl.acm.org/citation.cfm?id=1281253

29. Radev, D., Otterbacher, J., Winkel, A., Blair-Goldensohn, S.: Newsinessence: summarizing online news topics. Communications of the ACM 48(10), 95–98 (2005), http://dl.acm.org/citation.cfm?id=1089111

30. Sadiku, J., Biba, M.: Automatic stemming of Albanian through a rule-based approach. Journal of International Research Publications: Language, Individuals and Society 6 (2012)

31. Shahaf, D., Guestrin, C.: Connecting the dots between news articles. In: Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 623–632. ACM (2010), http://dl.acm.org/citation.cfm?id=1835884

32. Skenduli, M., Biba, M.: A named entity recognition approach for Albanian. In: Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on. pp. 1532–1537 (Aug 2013)

33. Subašic, I., Berendt, B.: Experience stories: a visual news search and summarization system. In: Machine Learning and Knowledge Discovery in Databases, pp. 619–623. Springer (2010), http://link.springer.com/chapter/10.1007/978-3-642-15939-8_47

34. Tintarev, N., Masthoff, J.: Similarity for news recommender systems. In: Proceedings of the AH'06 Workshop on Recommender Systems and Intelligent User Interfaces. Citeseer (2006)

35. Yang, Y., Pierce, T., Archibald, B.T., Carbonell, J.G., Brown, R.D., Liu, X.: Learning approaches for detecting and tracking news events. IEEE Intelligent Systems 14(4), 32–43 (1999), http://www.computer.org/csdl/mags/ex/1999/04/x4032.pdf

Knowledge Representation, Languages, Tools, and Frameworks in Medical Domain

Fesal Baxhaku[1], George Eleftherakis[2], and Eleni Vasilaki[3]

[1] South-East European Research Centre, University of Sheffield, Thessaloniki, Greece
[2] CITY College - International Faculty of the University of Sheffield, Thessaloniki, Greece
[3] Department of Computer Science, University of Sheffield, Sheffield, United Kingdom
febaxhaku@seerc.org, eleftherakis@city.academic.gr, e.vasilaki@sheffield.ac.uk

Abstract. Following an enormous growth of Internet users during the last decades, the new trend is to interconnect personal devices that typically feature a significant number of sensors. The availability of this vast amount of data brings unique opportunities and challenges: Exploiting these data real-time in an intelligent way will enable us to built sophisticated applications that can offer a better quality of life for the Internet users. In order to achieve this goal, we need to guarantee interoperability and smooth integration of the data provided by the sensors and at the same time to enable machines to interpret autonomously this primitive information, and produce higher-level knowledge. This paper provides a thorough literature review of technologies that would assist tackling these challenges, with specific focus on applications in the healthcare domain. We first present the state-of-the-art in knowledge representation. We further discuss ontologies as a means for enabling semantic annotation of sensor data and for supporting semantic interoperability between devices that would facilitate machine interpretation and potentially learning. In particular, we elaborate on medical ontologies representing knowledge in the healthcare domain. Finally, we discuss technologies that enable ontology-empowered sophisticated applications with the view of developing architectures that will facilitate non-centralized solutions to emerge.

Keywords: Knowledge Representation, Semantics, Ontologies, Sensor Data Annotation

1 Introduction

Digital technologies are able to transform public and health care services for citizens though exploiting the vast amount of data publicly available and improving the use of resources. In particular, the ability of interconnecting sensory devices via the internet provides us the means of collecting distributed information that can be used to better monitor the environment. A recent prediction [40] indicates that until 2020 the number of connected devices will run up to 100 billions, which is ten times higher compared to the whole population.

Healthcare is one specific area where exploiting distributed sensory information can bring vast improvements. There are already instances of "smart hospitals"[43] that allow to continuously monitor both patients and staff. Typically, though, such designs are developed as apart of the internal, closed infrastructure of the hospital, while our vision is to provide an "open architecture" which offers a distributed, non centralised solution allowing external users to join the network by providing own sensory data.

In order to achieve this vision, we need to enable the interconnection of a variety of sensors in the network, collecting data described consistently across all modalities. For this purpose, an appropriate mark-up language should be used to annotate useful information within the network. A machine, however, will not be able to interpret this annotation without the use of an ontology, which will provide the context for interpreting the collected data. Such an ontology will allow the design of autonomous, smarter applications that can make sense of the measurements and provide higher level interpretation of the data.

In this paper, we present technologies which support semantic annotation of stream sensor data in the healthcare domain. Due to the high cost and narrow services in healthcare, the demand for sophisticated services will increase with the ageing population [7] and the advancements in sensor technology. Taking advantage of the technological progress, we are now able to build autonomous and pro-active healthcare solutions [7].

The structure of the paper is as follows: In section 2, we present the concept of knowledge representation and its application areas. We particularly focus on ontologies as a mean for enriching data with semantics and we present the most prevalent semantic web languages. Section 3 presents data annotation tools and section 4 discusses findings and future directions.

2 Knowledge Representation

Most of the information available, exists in documents written in natural languages, while the highest percentage of data are included in relational databases. Due to the advancement in sensor technology, the amount of data also increases, hence we seek for an effective and intelligent way to transform raw data into knowledge and service repositories. Representing information in machine interpretable format is significant in Artificial Intelligence and Computer Science in general.

Knowledge Representation (KR) is the field of Artificial Intelligence that deals with the annotation a set of sentences using formal symbols, appropriate for agents to apprehend relevant information [12]. According to [23], KR aims to design computer systems such that machine interpretations of the world are similar to human perception. This is achieved by using symbols which serve as surrogate for domain entities, events, relations.

The main purpose of KR as pointed out by [10] is to provide Intelligent Systems (IS) with information in a specific domain with the view to enable efficient process of data. Ontology is a form of Knowledge Representation appropriate

for modelling notions about the world such as entities, activities, relations and goals [26]. In what follows, we will further discuss Ontology and in particular within the context of healthcare applications.

2.1 The Science of meaning and Semantics Web

Semantic [4] refers to the study of meaning, or interpretation, of a word, or a sentence. Computers require a semantic representation in forms of symbols in order to understand the meaning of a word or a statement [38].

Semantic technologies, therefore, enable machine understanding and interpretation. The current web, though one of the most successful engineering artefact created so far [21], has perhaps not exploited sufficiently semantic technologies. The semantic web is an extension of the current web with a meaningful structure that allows humans and agents to cooperate with each other and accomplish sophisticated tasks [11].
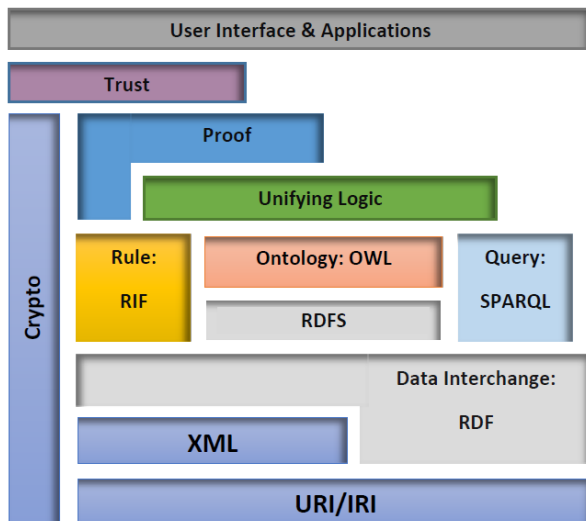
Fig. 1. Semantic Web Architecture adopted from [18].

The technologies presented in Fig. 1 transform the web to a suitable platform for developing applications where the data can be shared and reused, with the view of enabling efficient knowledge discovery [42]. While the current web is dedicated for humans understanding, the semantic web focuses human-computer cooperation, using same principles and technologies of the web [21]. In the semantic web, meaning is a shared resource which involves Ontology in order to enable knowledge discovery.

---

[4] Semantic according to: http://www.thefreedictionary.com/semantic

2.2 Ontologies

Ontology, as a branch of Philosophy, is the term used for describing things that exist [26, 14] or existence of being in the world [39]. In Computer Science, Ontology is used for knowledge representation [26], natural language processing and knowledge engineering [39]. According to [39], an ontology is a formal, explicit specification of a shared conceptualisation.

Ontologies have improve the interoperability between agents by enriching semantically digital content and providing machine understandable content [44]. In other words, Ontology is used to share common understanding of the structure of information for a particular domain by agents and people, reuse of those information, analyse and finally provide agent interpretation.

Ontologies provide concepts which represent a set of classes for entities in a domain of interest, relations by relating classes, and instances that presents the objects [23, 30]. In comparison to object oriented modelling, ontologies provide domain knowledge where automated reasoning is possible, also reacher formal semantics like encoding the complex information in a specific domain [23]. In addition, ontology allow different level of relationships among classes, or numerous inheritances of classes and properties, that is not the case in object oriented modelling. For instannce, in object oriented we can have a hierarchical relation up to sublcasses while in ontologies we can have relation even in property level.

Application areas of ontologies. Information-retrievals systems, digital libraries, and internet search engines make use of ontologies for searching and organising information [26]. Ontologies are useful in different application areas [23]. Some of these applications are listed below:

- **Information integration:** Ontologies assist in the integration of heterogeneous sources of data under a schema.
- **Information retrieval:** Ontologies enhance the quality of the result and answer user queries with high precision, via semantic matching of key terms with ontological concepts and their relations.
- **Semantically enhanced content management:** Ontologies facilitate the annotation process of data by providing domain specific vocabulary and thus enhancing the level of expressiveness as well as enabling machine interpretability.
- **Knowledge management and community portals:** Ontologies assist in connecting technical systems that search and exchange community knowledge.
- **Expert Systems:** In sensitive areas such as medicine Ontologies could help emulating experts, by providing a way of answering sophisticated questions.

2.3 Ontologies in Medical domain

Over the years, ontologies are developed aiming to specifically address problems in the medical domain, for instance Protein Ontology, WordNet Semantic

Lexicon, Foundational Model of Anatomy (FMA Ontology), Gene Ontology, Medical Subject Heading (MeSH), Current Procedural Terminology (CPT) and many others. Up to today, there is no standard ontology in the medical domain, despite numerous attempts.

We, here present two of these ontologies: Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) and Unified Medical Language System (UMLS).

Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT). SNOMED CT is a large, comprehensive clinical terminology, developed by an international body (IHTSDO) with eighteen country members in 2012 [6]. It is a recommendation reference for clinical information systems to United States, United Kingdom, New Zealand, Canada and Australia [31, 36]. In [27] indicates that SNOMED CT provides, around 400 000 pre-defined concepts, with over 300 000 active and over a million relationship between them. It is suitable for encoding clinical terms where many applications are developed, especially in supporting Electronic Health Record systems.

SNOMED CT ontology is organized in hierarchy where the top node consist of root concepts and each other node is subtype of root node. In the top level nodes there are concepts which are intended to represent real world entities, events and processes, concepts that represent clinical terms. Top level nodes defines concepts for: clinical findings, procedure, observable entity, body structure, organism, substance, pharmaceutical/biological product, specimen, special concept, snomed ct model component, event, environment or geographical location, physical force, social context, situation with explicit context, staging and scales, physical object, qualifier value, and record artifact [36].

According to IHTSDO technical guide three different implementation areas of SNOMED CT are: clinical records, knowledge representation and aggregation and analysis [36]. It can be used in knowledge management, decision support systems, facilitate performance of patients data and analyse them, improves clinical efficiency, supporting cost-effective delivery care etc. [36].

Unified Medical Language System (UMLS). Is a project by National Library of Medicine for enhancing and facilitating the development of systems in biomedicine by providing meaningful resources. Its knowledge resource is dedicated to work as multipurpose vocabulary that can be used for various systems that performs a range of functions starting from patients record, scientific literature, public health data, and guidelines [4].

UMLS consist of three major components: Metathesaurus, Semantic Network and the Specialist lexicon [4]. Metathesaurus is a vocabulary which consists of concepts from different vocabulary systems, thesauri, code sets and other related terminologies in biomedicine and health related concepts. Semantic Network provides a way of categorizing concepts in UMLS and provides relationship between concepts. It consists of 135 semantic types which organize concepts and 54 relationships for semantic relations between semantic types. Specialist lexicon refers

to lexical terms in biomedical domain, English vocabulary and other terms found in Metathesaurus which describe the syntactic, morphological and orthographic information which are useful by Natural Language Processing systems.

Currently, Metathesaurus counts 150 different code/vocabulary systems from different medical domain including Mesh, LOINC, SNOMED CT, RxNorm where some of them comes with different releases [2]. Metathesaurus consist of concepts that make the organization structure. The upper ontology of UMLS in semantic network consist of Entities and Events. Below that, there are major categories which describe organisms, anatomical structures, biologic function, chemicals, and physical objects [3]. Each semantic type has a textual description which is displayed in the top of hierarchy. In addition to taxonomies there are five kind of relationships between semantic types: physical relations, spatially, temporally, functionally, and conceptually. Each semantic type consist of unique identifier, a tree number, a definition, its immediate parent and children.

2.4 Semantic Web Languages

In our specific application, a key concept is the language describing sensor data, which would enable automatic machine processing. Defining an ontology language common standard is of great importance for Semantic Web community, with high impact for both industry and academy [23]. In fact, the community has been developing a set of core technologies where some of them are de facto standards for the semantic web [33]. Standardizing an ontology, however, requires taking into consideration different aspects, including expressivity and computational capability of the language, web-related features and XML serialization syntax or usability of add-ons for enriching with meta data [23]. Besides this, choosing a language for ontology development depends on specific requirements. In what follows, we mention a number of excising ontology languages:

- **Ontolingua** is a language used by the Ontolingua server [20] appropriate for building ontologies in a canonical format used for knowledge representation systems, reasoning and share among different systems [24].
- **CycL** is a formal language whose aim is to empower AI application to perform human reasoning [1]. CycL originally was developed as frame based language while later it becomes more declarative language [1].
- **Operational Conceptual Modelling Language (OCML)** is a frame-based language for knowledge modelling and problem solving method. In addition it allows specification and operationalization of functions, relations, rules, classes and instances [35, 34].
- **Frame Logic (F-Logic)** is a language which combines features of object-oriented and frame-based languages. The terminology of F-Logic comes from object-oriented phrasing, but with feature combination from frame-based languages [29].
- **LOOM** is a knowledge representation language for developing intelligent application where reasoning is required. The declarative knowledge consist of definitions, rules, facts, and default rules [5].

- **Ontology Inference Layer (OIL)** is a Web-based representation and inference layer for ontologies. It is based on Description Logic for reasoning services and frame-based languages for semantic modelling [22].
- **DAML+OIL** a semantic markup language for web resources which combines features of DARPA agent markup language and OIL. It is built on previous W3C standards, RDF and RDFS and its knowledge base is a collection of RDF triples [25].
- **Extensible Markup Language (XML)** is a markup language designed with the aim of separating web data from web data presentation. It is interoperable with html and dedicated for easy implementation. It was developed by XML working group under W3C 1996. One of the drawbacks that is mention is the lack of semantics, even though it is a web standard for representing information [13].
- **RDF** is a standard for representing data on the Web. Data are represented as a triple where each triple contains a subject node, predicate and object node while the structure of the web is extended by using URI to name the relationship between things [16]. RDF is easy to implement and provides a formal semantics that can be reason about the meaning of an RDF expression. In addition, XML serialization form (RDF-syntax) can be used for exchanging data between applications [33]. Because of simplicity in data representation, in RDF is not possible to relate concepts, relation between concepts because it is limited to low level expressiveness. To overcome these limitation RDF Schema and OWL are proposed.
- **RDF Schema (RDFS)** is an extension of RDF vocabulary and a specification language for describing data properties based on RDF [17]. RDFS is simply an RDF extension with additional statements whereas the combination of both technologies form a simple ontology language [21]. RDFS introduce the notion of classes and properties which make it similar to object oriented languages such as Java, but instead of writing classes in terms of properties, RDFS describes properties in terms of the classes of resource to which they apply [17].
- **Web Ontology Language (OWL)** is a standard for knowledge representation on the web [17]. It is an extension of the RDF(s) vocabulary, based on the DAML-OIL language, which is an older semantic markup language for Web resources. With the increasing need for adding more features in OWL ontologies, and making the Web content more accessible to machines, while increasing the reasoning power, [37, 8]. In 2007, the W3C Web Ontology Working Group releases a new standard OWL 2. OWL like RDF can be written in different syntactic forms, whereas the primary syntax in OWL 2 is RDF/XML. However, alternative syntaxes such as alternative RDF serialization, XML serialization or other different format like Manchester can be used.

The Ontology Web Language (OWL) is widely adopted by the research community [35]. Below we present the structure of OWL 2. In this schema, the ellipse representw the ontology as an abstract structure or RDF graph, the syntax for

serialization or exchange is represented on the top while the semantics that specifies the OWL 2 meaning [37] is represented at the bottom.
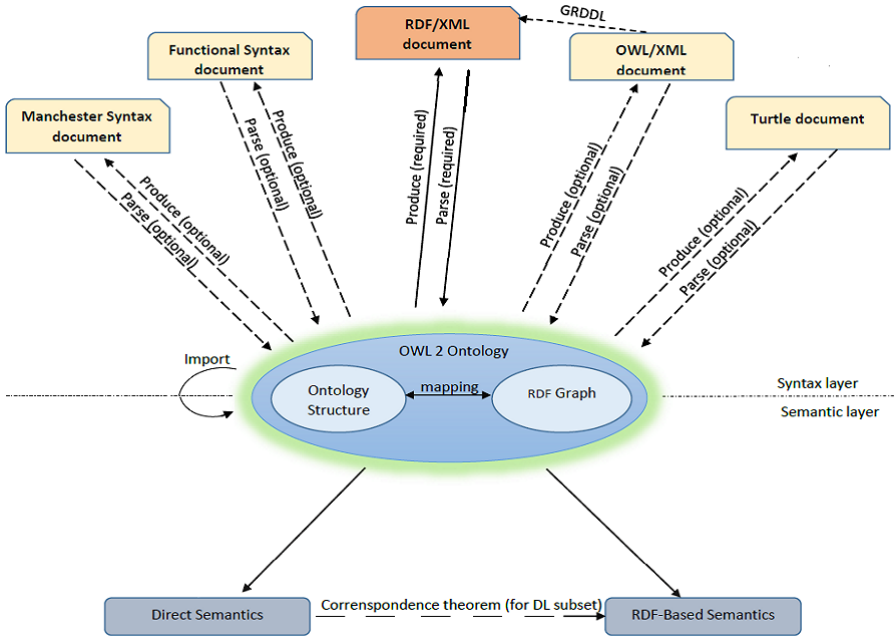
Fig. 2. OWL 2 Architecture adopted from [37].

Depending on the needs of knowledge model expressivity, three sublanguages of OWL are distinguished [37] :

– **OWL Lite** provides simple features, and designed for those who need to define classification hierarchies for describing ontologies
– **OWL DL** is a language which support existing description logic languages and provides computational properties for reasoning, it has a finite time to finish computations and entailment knowledge is granted to be computed.
– **OWL Full** it adds more expressive power to OWL DL as well as syntactic freedom of RDF but the entailment knowledge is not granted to be finished and is not decidable.

OWL Lite implements just part of OWL language and can be considered as sublanguage of OWL DL, while OWL Full and OWL DL has the same construct but they differ in the restrictions they provide. Because of the expressive power of OWL DL and decidability of this language, it is widely accepted and many applications are made using this language. OWL Full is less accepted as it is not a decidable language.

In Medicine one of the most widely used for building ontologies is Open Biomedical Ontologies (OBO) format, because of its simplicity and an easy understanding structure it provides. OWL 2 provides an easy translation way from OBO to OWL.

3 Data Annotation Tools

A large number of tools have been created with the purpose of annotating web and textual documents. This section presents selective tools and frameworks for enriching textual data with semantic annotation.

Tools and Frameworks. **MeatAnnot** is a system for enhancing medical documents with semantic annotations. UMLS vocabulary serves as reference for detecting important terms and relations within textual documents [28]. Natural Language Processing tools such as Rasp, TreeTagger and Gate serve for the purpose of term extraction within a sentence and document as a whole. Three phases exist in semantic annotation: relation detection, term extraction and annotation generation. In addition, semantic annotations are stored as RDF triples.

Cream (CREAting Metadata) is a framework for annotating web pages by looking the context of definitions of classes from ontologies and utilise its annotation format [32]. Furthermore, Cream provides a way for annotating data coming from databases [41]. In addition, technologies such as RDF, OWL and XPointer [5] are forms of representing annotations in text.

Annotea is another framework, which make possible annotation of web documents. Unlike CREAM, annotations are inadequate for databases and limited to XML and HTML documents [41]. In addition, RDF is a model for storing annotation in web documents.

There are many other tools that aim to enriching web documents with metadata, some of which require manual annotation while others are partially automated. A summary of selected annotation tools is presented in the Table 1.

---

[5] XPointer w3c recommendation for identifying fragments of URI resources

<p align="center">Table 1. Some annotation tools presented in [41].</p>

Tool Name	Automation	Type of analysis	Learning technique
Amay	No		
Mangrove	No		
Ontomat	Yes	Pankov, Amilcare	Supervised Learning
M-Ontomat-Annotizer	Yes	Extraction of spatial descriptors	Genetic algorithms
SMORE	Yes	Screen Scraper	No
Open Ontology Forge	Yes	String Matching	No
COHSE annotator	Yes	Ontology string matching	No
Lixto	Yes	Wrappers	No
MnM	Yes	Pos tagging, Named Entity Recognition	Supervised learning
Melita	Yes	String matching, POS tagging, Named Entity Recognition	Supervised learning
Parmenides	Yes	Text minning with constraints	Unsupervised learning

4 Discussion

There is a need for efficient methods to transform data into smart or intelligent data. Different layers of knowledge is required to transform data into smart data (Fig. 3) and provide useful service to built sophisticated applications that can offer a better quality of life for the users. For instance, by combining sensor data from a patient that involve high blood pressure and high cholesterol, an intelligent system may conclude that patient has a high heart attack risk. Such an application would be very much desirable: The Center for Technology and Aging in their report estimates a reduce cost of $ 200 billion dollar in next 25 years with inclusion of remote monitoring systems for congestive heart failure, chronic obstructive pulmonary disease, and chronic wounds or skin ulcers [19]. Major companies like Humana, AT&T, Healthsense, and Qualcomm are addressing their potential for remote monitoring systems [19]. "Having reliable data about a patient's condition is essential to responding to changes at the earliest possible time and preventing deterioration of their condition or even death" [19].

Semantic technologies play crucial role in data integration, semantic interoperability and knowledge discovery. Ontologies could help to automate the process of semantic annotation of both existing medical documents and sensor data and transform in useful information. Standard Ontologies in the medical domain will enable the design of sophisticated applications with reasoning ambitions. Out of many ontologies, SNOMED CT seems to be widely accepted for clinical terms, whereas UMLS holds different Ontologies including those with analogue purpose. The reason for adapting SNOMED CT is probably the standard they follow in naming, while UMLS provide list of similar terms from different ontologies which makes it hard to decide which one to adopt. Semantic ontologies and
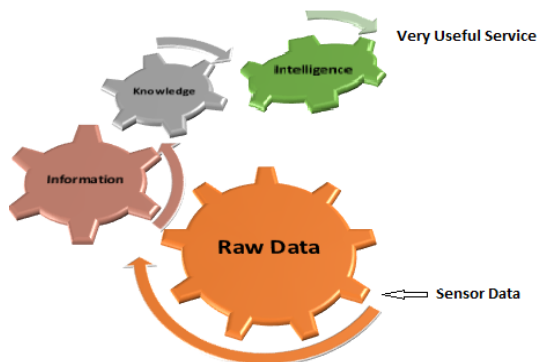
Fig. 3. Process of data transformation from non-sense information to useful services adopted from [15].

machine-readable codification could run over problems such as misunderstanding or misinterpretation and cross-referencing between relevant data existing in different systems [9]. Nevertheless, for those interested in exploring genetics, Gene Ontology is a project with the aim of standardizing the representation of gene and its attributes for species.

Regarding languages which support machine processing and uphold semantic description of objects, attributes and relations, two standards have been identified: RDF and OWL. Until now, XML was extensively accepted but it has its limitations in supporting semantic description. RDF supports xml for structuring information, and has more expressive power in providing semantics.

In summary, we have identified key technologies for supporting semantic stream data annotation. Using such technologies will help us to enrich data streams with knowledge and transform them into smart data. Our future work aims to automate as much as possible the semantic annotation process of continuous data generation from sensors who monitor people with chronic diseases.

References

1. CYCL Syntax, `http://www.cyc.com/cyc/cycl/syntax,` Last accessed: march, 2014
2. UMLS Reference Manual, `http://www.ncbi.nlm.nih.gov/books/NBK9684/`, Last accessed: april, 2014
3. UMLS Semantic Network, `https://www.nlm.nih.gov/pubs/factsheets/umlssemn.html`, Last accessed: april, 2014
4. U.S National Library of Medicine, UMLS, `https://www.nlm.nih.gov/research/umls/about_umls.html3`, Last accessed: april, 2014
5. Loom Project (2007), `http://www.isi.edu/isd/LOOM/LOOM-HOME.html`, Last accessed: march, 2014

6. International Health Terminology Terminology Standards Development Organization, SNOMED (2012), `http://www.ihtsdo.org/snomed-ct/`, Last accessed: april, 2014
7. Acampora, G., Cook, D.J., Rashidi, P., Vasilakos, A.V.: A Survey on Ambient Intelligence in Health Care. Proceedings of the IEEE. Institute of Electrical and Electronics Engineers 101(12), 2470–2494 (Dec 2013)
8. Andrea, M., Francesco, R., Maurizio, T., Salvatore, M.: Formalizing Knowledge by Ontologies : OWL and KIF
9. Bandayopadhyay, D., Sen, J.: Internet of Things - Applications and Challenges in Technology and Standardization
10. Berner, S.E., Hannah, J.K., Ball, J.M.: Clinical Decision Support Systems (2007)
11. Berners-lee, T., Hendler, J.: The Semantic Web. Scientific American 21(May 17), 29–37 (2001)
12. Brachman, J.R., Levesque, J.H.: Knowledge Representation and Reasoning (2003)
13. Bray, T., Paoli, J., Sperberg-McQueen, C.M., Maler, E., Yergeau, F.: Extensible markup language (xml) 1.0 (2004)
14. Chandrasekaran, B., Josephson, J., Benjamins, V.: What are ontologies, and why do we need them? IEEE Intelligent Systems 14(1), 20–26 (Jan 1999)
15. Chen, Y.: Challenges and Opportunities of Internet of Things pp. 383–388 (2012)
16. Dave, B. , and Brian M.: RDF/XML syntax specification (revised). W3C recommendation 10 (2004)
17. Dean, M., e.a.: OWL Web Ontology Language 1.0 Reference, W3C Working Draft 29 July 2002 (2002)
18. Desjardins, D.: Decentralized Extensibility, `http://www.w3.org/2011/Talks/0223-cshals-egp/#\%2811\%29`, Last accessed: may, 2014
19. Desjardins, D.: Remote Monitoring Exploring New Territories (May 2014), `http://healthsense.com/index.php/press-releases?id=301`, Last accessed: april, 2014
20. Farquhar, A., Fikes, R., Rice, J.: The Ontolingua Server: a tool for collaborative ontology construction. International Journal of Human-Computer Studies 46(6), 707–727 (Jun 1997)
21. Fensel, D., Hendler, A.J., Domingue, J.: Introduction to the Semantic Web Technologies pp. 1–28 (2012)
22. Fensel, D., Horrocks, I., Harmelen, V.F., Decker, S., Erdmann, M., Klein, M.: OIL in a Nutshell pp. 1–12
23. Grimm, S., Hitzler, P., Abecker, A.: Knowledge Representation and Ontologies Logic, Ontologies and Semantic Web Languages
24. Gruber, R.T.: Ontolingua: A mechanism to support portable ontologies (1992)
25. Harmelen, v.F., Patel-Schneider, F.P., Horrocks, I.: Reference description of the daml+ oil (march 2001) ontology markup language. DAML+ OIL Document (2001)
26. Jurisica, I., Mylopoulos, J., Yu, E.: Ontologies for Knowledge Management: An Information Systems Perspective. Knowledge and Information Systems 6(4), 380–401 (Mar 2004)
27. Kate, R.J.: Towards Converting Clinical Phrases into SNOMED CT Expressions. Biomedical Informatics Insights 6(Suppl 1), 29–37 (Jan 2013)
28. Khelef, K., Dieng-Knutz, R., Barbry, P.: An Ontology-based Approach to Support Text Mining and Information Retrieval in the Biological Domain. Universal Computer Science 13(12), 1881–1907 (2007)
29. Kifer, M., Lausen, G., Wu, J.: Logical foundations of object-oriented and frame-based languages. Journal of the ACM 42(4), 741–843 (Jul 1995)

30. Lambrix, P.: Towards a Semantic Web for Bioinformatics using Ontology-based Annotation (2005)
31. Lee, D., Cornet, R., Lau, F., de Keizer, N.: A survey of SNOMED CT implementations. Journal of biomedical informatics 46(1), 87–96 (Feb 2013)
32. Macário, C.G.N., Medeiros, C.B.: A framework for semantic annotation of geospatial data for agriculture. Int. J. Metadata, Semantics and Ontologies (2009)
33. Machado, M.C., Rebholz-Schuhmann, D., Freitas, T., Couto, M.F.: The semantic web in translational medicine: current applications and future directions. Briefings in bioinformatics (11 2013)
34. Motta, E.: An Overview of the OCML Modelling Language (1998)
35. Nguyen, V.: Ontologies and Information Systems : A Literature Survey (2011)
36. Organization, I.H.T.S.D.: Snomed CT Technical Implementation Guide
37. OWL Working Group: Owl 2 web ontology language: document overview. w3c recommendation (27 october 2009) (2012)
38. Peter, J., Wine, M.: Medical Semantics, Ontologies, Open Solutions and EHR Systems. Virtual Medical Worlds
39. Studer, R., Benjamins, V.R., Fensel, D.: Knowledge Engineering : Principles and methods 25, 161–197 (1998)
40. Sundmaeker, H., Saint-exupéry, A.: Vision and Challenges for Realising the Internet of Things. No. March (2010)
41. Uren, V., e.a.: Semantic annotation for knowledge management: Requirements and a survey of the state of the art. Web Semantics: Science, Services and Agents on the World Wide Web 4(1), 14–28 (Jan 2006)
42. Vanitha, K., Yasudha, K., Soujanya, K.N.: The Development Process of the Semantic Web and Web Ontology 2(7) (2011)
43. Yao, W., Chu, C., Li, Z.: Leveraging complex event processing for smart hospitals using RFID. Journal of Network and Computer Applications 34(3), 799–810 (May 2011)
44. Zhou, L.: Ontology learning: state of the art and open issues. Information Technology and Management 8(3), 241–252 (Mar 2007)

Towards a Wireless Monitoring System in Developing Regions – The Case of Kosovo

Adelina Basholli[1], Thomas Lagkas[2], Peter A. Bath[3], and George Eleftherakis[2]

[1]South-East European Research Centre, University of Sheffield, Thessaloniki, Greece
[2]CITY College- International Faculty of the University of Sheffield, Thessaloniki, Greece
[3]Information School, University of Sheffield, Sheffield, U.K.
adbasholli@seerc.org, {tlagkas, eleftherakis}@city.academic.gr, p.a.bath@sheffield.ac.uk

Abstract. Provision of health care services through the means of Wireless Sensor Networks (WSNs) is currently an area under intensive research. Taking into account people living in developing countries, the application of WSN architecture would help to improve their lives and provide a helpful and sophisticated health provision service. One of the main considerations while analyzing developing countries is affordability. The proposed system architecture needs to suite the general population's economic situation. With the aim of proposing a WSN system architecture for a newly-created Balkan country, Kosovo, we have analyzed the current situation of the health system and telecommunication infrastructure. WSN architecture requires the appropriate telecommunication equipment and services, in order to allow efficient exchange of information sensed by sensors, as well as feedback from medical specialists. The current telecommunication infrastructure is quite promising for the future implementation of wireless-based solutions. In this paper we present data and analyses on patients' needs for health services facilitated by WSNs.

Keywords: Wireless Sensor Networks, Wireless Health Monitoring System, Developing Countries

1 Introduction

Healthcare applications nowadays are seen as an emerging need when considering the ageing population, chronic diseases, acute care, and the need for early diagnosis. On the other hand, the wide availability of wireless networks, their increased bandwidth, and the advances in the miniaturization of a sensor's dimensions provide many possibilities to research and develop a Wireless Sensor Network (WSN) architecture that may help the health system.

Wireless Sensor Networks present architectures of sensor nodes which are able to sense biological information when placed near to a patient's body. The sensed information can be transmitted and further processed and analyzed by a medical specialist. In this way, the patient's condition can be monitored continuously and the health provider can have a detailed report of the patient data. Hence, the application of wearable health-monitoring devices, which are part of the Wireless Sensor

Network's system architecture, can be integrated into a person's clothes. As follows, they can monitor a diagnostic procedure continuously and supervise a chronic patient's condition or different surgical procedures.

Remote medical advice based on the parameters received through wireless connection from a sensor, will make easier medical care for patients and doctors in general. When considering developing regions where the doctor's availability and presence is limited, these applications may provide helpful systems for patients to gain better access to health services.

The Republic of Kosovo, as a developing country, provides a suitable case study for our research. As a newly created country, and one that recently experienced war, Kosovo's health system needs improvements, especially considering that it is handling a large number of patients. The current conditions and infrastructure are promising and they are improving; however, continuous and regular health provision still remains below the desired levels. Moreover, the application of a WSN needs a good network infrastructure. This is another advantage that Kosovo exhibits. There have been considerable investment and provision of new telecommunication services that can facilitate the implementation of a WSN architecture.

In this paper we present some related approaches which describe previously proposed or implemented WSN architectures, which are mostly specific for developing regions. In section three, a case study is presented, continuing with some demographic data. Section four includes a description of the Kosovo health system and mainly the phases that it has followed. We then present and discuss the current network infrastructure in Kosovo in sections five and six. The final section concludes this paper.

2   Related Approaches

The application of Wireless Sensor Networks in developing regions is considered a helpful mechanism for providing continuous healthcare services. The use of Wireless Sensor Networks has many advantages that can help people living in developing regions, such as [1]:

- Continuous monitoring of living parameters
- Access to specialists
- Detailed data about patient
- Improved healthcare outcomes
- Cost savings

In order to have the above-mentioned facilities, an implemented architecture is needed. Considering the developing countries, this architecture should suite their particular needs and conditions. Researchers in [2] found that WSN applications for developing countries should have continuous connectivity, should present low-cost products, should be user friendly for ease of use, and be power resilient.

Researchers at the University of Texas, Medical Branch, concluded that the use of smart phones, whose performance capabilities are enhancing day-by-day, and the small sizes of sensors will enable the use of wearable/portable embedded architectures

that will facilitate the patient's access to a medical specialist or a healthcare centre. Further to this, researchers in [3] proposed a solution for the issue of accessing medical services for people who live in rural areas in developing countries. This solution consists of a Medical Expert System operating in a semi-automatic and mobile hospital way, which enables the data collection of all sensed information from different Body Area Networks placed in the patient's body. In this way, a person will assist in gathering the data from the people who visit the healthcare centre. In cases of abnormal conditions, the assistant will contact the doctor directly. In this way, after finishing one village, the machine is moved into another village and so on, visiting rural areas periodically and providing healthcare services. However, this is not an efficient solution, as it does not provide continuous health care.

In practice, some similar systems that have aimed to facilitate patients' lives have already been implemented. A similar application of Wireless Sensor Network architecture is presented in [4]. This system architecture introduces the appliance of sensors and actuators in a 'hospital centre' (improvised) to monitor the heartbeat, temperature and motion of a pregnant woman. The authors have implemented, tested, and compared the proposed model architecture, which consists of multiple patients and multiple functional layers. Based on the gathered results, this application promises the reduction of energy consumption which is related to network lifetime, and speed of data transmission. In this way, the whole system flexibility to ensure patient better life will be enhanced. Furthermore, this architecture realizes the patient-physician communication through the use of a GSM modem that enables sending SMS messages in emergency situations. However, this system was tested in an improvised hospital centre without any obstacles such as walls or doors. In addition, the usage of Time Division Multiple Access assumes only a fixed number of nodes (because of the allocation of time slots). Consequently, if the number of nodes in the Wireless Sensor Network architecture needs to change, it would be difficult for a TDMA protocol to adapt accordingly to the number of additional nodes, because of the frame length and time assignments. A very similar application is developed for smart phones and is used to monitor fetal heart rhyme. This application is commercial and it is called AirStrip OB [5].

Cardiac diseases are considered one the main causes of death of people living in developed countries. Researchers in [6] presented a mobile health monitoring application which is capable of analyzing the electrocardiography (ECG) waveforms. The ECG monitoring mobile application presented in their paper has a receiver that receives the signal transmitted from the source, analyzes it (it receives an analog signal, converts it to digital and then this signal is processed through filters and other mechanisms to remove the noise and other inconstancies). At the end the data are presented through five different screen activities. In this way, people who need special attention and care are notified for possible problems that they need to address. However, as the authors suggest, this application needs further enhancement in signal processing techniques and sensor application mechanisms, because of the sensitivity of the application.

The famous telecommunication company in China, Qualcomm, and Life Care Networks present a system based on 3G wireless workstations to monitor patients with cardiovascular diseases. The initiative is called Wireless Heart Health [7]. This system architecture consists of smartphones that are able to record electrocardiogram

data and signals through the means of build-in ECG sensors. Patients have to carry this remote monitoring mobile application which is able to automatically send patient data over the 3G network to a health specialist. The call centre provides 24-hour monitoring and gives feedback to the possible alarms by SMS messages or phone calls. However, in critical cases the specialists may suggest hospitalization for further treatment or testing.

In general, the technology is enhanced day-by-day, providing in this way different solutions to healthcare issues. Therefore, taking also into account the developments in healthcare providing methods and medicaments, the physicians should continuously stay informed and in touch with these enhancements. On this basis, the author in [8] suggests the creation of an online platform that will enable physicians to stay up-to-date with the latest medical information and corresponding applications. Similar platforms already exist, such as: mPowering Frontline Health Workers, or a private library in a healthcare centre in South Africa.

Taking into account already proposed and implemented WSN architectures, and also considering the benefits that they provide to a developing country, in the following sections we present a detailed analysis of the current need and infrastructure for our case study, Kosovo.

3   Presenting our Case Study- Kosovo

Kosovo is located in the Balkan Peninsula of Southeastern Europe. The total number of resident population is 1,815,606 [9] and 61% of the population lives in rural areas. The General Statistics for Kosovo are as follows:

- − Capital city: Prishtina
- − Geographic Area: 10,908 km^2
- − Border countries: Albania, Montenegro, Serbia, Macedonia
- − Religions: Muslims over 90%, Orthodox Christians 6% and Roman Catholics are 3% and close to 1% Protestant.
- − GDP: GDP (International Monetary Fund (IMF), 2011 estimate): $6.5 billion. GDP (Kosovo Ministry of Economy and Finance (MEF), 2010 estimate): $7.5 billion [10]

The country is characterized by a young population, such that 28% of the population is less than 15 years old, and half of the total population is less than 28.2 years old [11]. Kosovo is one of the countries with a high number of births. However, the number of the population is expected to decrease during the coming years. This is estimated in a study presented in [11].
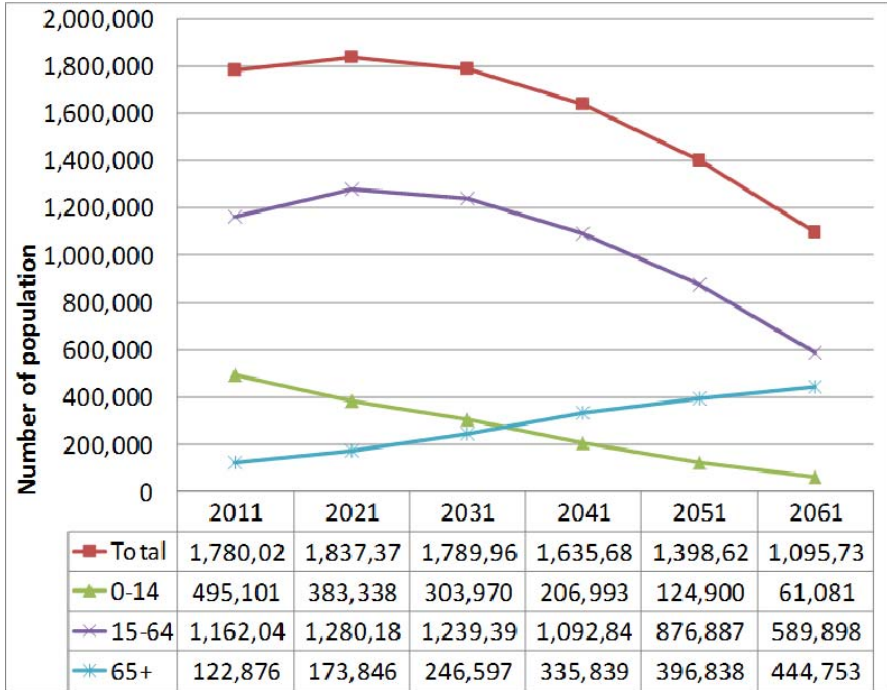
Fig. 1. Kosovo Population based on age between years 2011- 2061 [4]

Figure 1 presents some estimated numbers of the population in the next coming years based on the current situation and the corresponding life expectancy parameters. Considering various factors, such as: female education, her employability, economic situation, etc; the number of births is expected to decrease further from 2.0 children per woman as it was in 2011, to 1.7 children per woman in year 2031 and 1.5 in year 2061 [11].

Besides the decrease in the number of the young population, Kosovo is expected to have an increased population of elderly people. Therefore, the life expectancy is going to increase, and mortality rates will decrease, by 8.5 years for an estimated period of 50 years [11].

4   Review of the Health System in Kosovo

Kosovo, as a country that has lately experienced war, is considered to be a developing country with a sensitive health system. In general, the challenges can be categorized mainly into two periods: the period before the war and the period after the year 1999, when the war ended. Before 1999, Kosovo had a massive, hierarchical, and centralized system inherited from the Yugoslavian government. This system had many problems, such as employees with low levels of preparation, low prestige, and very low salaries that were not motivating. These were caused by yearly

discrimination from the Yugoslavian dominion. Therefore, Kosovo doctors lacked the professional experience, practice, usage of new methods and treatments based on international standards [12, 13]. Furthermore, after a decade without investments, the infrastructure was old and consumed. This system became worse during the political and economic conflicts. Approximately 90% [14] of clinical and health institutions were damaged during the war. Moreover, other essential needs such as water and electricity, and the public service infrastructure had a very negative impact on the health sector of Kosovo [15].

After the conflict of 1999, there were very limited capabilities for the health institutions to offer services requested by patients. Besides the central health centre, different reports [14] state that the health giving institutions in rural areas were lacking medical staff and equipment [16]. This was due to several factors:

- Access to health services were non-consistent
- Lack of medical staff
- Service Efficiency
- Double and inefficient work caused by a decentralized health system

In order to handle huge requests after the conflict, the health systems in Kosovo hired uncontrolled personnel [14]. Following this, more than 400 donors from all over the world associated and helped Kosovo to reconstruct its health system [15]. Help came in different sectors, such as basic health services, investments in public infrastructure, further education, and professional trainings for doctors [12].

However, besides the critical moments that the Kosovo health system has passed, in a research study conducted recently (November 2013) [14], the patient satisfaction with the health care system seemed relatively good. Even though we should consider that the actual real situation in clinical centres depends also on the technical conditions, for example, a doctor has five or six patients in one room, with uncomfortable and unsuitable conditions for offering proper care and services to each patient.

The patient satisfaction with the health offered services and care in health institutions in Kosovo is presented in Figure 2. These data were gathered in a study conducted in November 2013 in different health care institutions, starting from: Family medicine centre, Hospitals, University Clinical Centre of Kosovo, Private health providing institutions, and Pharmacies.
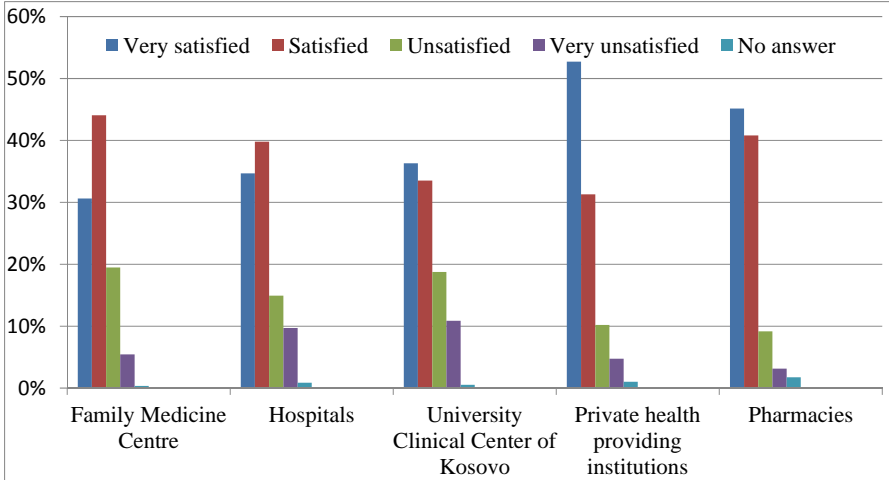
Fig.2 Patient satisfaction with health offering services in health institutions in Kosovo [14]

Based on gathered results in [14], more than 70% of the investigated people were satisfied with the offered health services in Kosovo health institutions.

4.1   Need for continuous public health- Personal Doctor

Based on the statistics drawn from the research presented in [14], 80% of around 1334 interviewees declared the need for continuous attention and care from the health institutions. However, based on the same report, more than 85% of them do not have a personal doctor to address their needs or anyone who can continuously monitor their condition. Only 15% of the interviewees confirmed that they do have a personal doctor to consult, visit, and take care of them. These results are closely related to the general education and knowledge level of patients about the benefits and need for a personal doctor. In this study, more than 85.20% of the interviewees answered negatively on the need for a family doctor to address emergency cases, for continuous monitoring, and for the provision of other medical services.

However, despite the lack of personal doctors, more than 85% of the interviewees were very satisfied with the process of health care services, and confirmed that the treatment given to them was successful. On the other hand, these patients agreed about the limited time that doctors and nurses can spend with them, due to the high number of patients and low number of health providing personnel [14].

5   Network Infrastructure in Kosovo

Following the 1999 conflict, Kosovo was also damaged in its network and communication infrastructure. Since then, the Kosovo market has experienced a remarkable change and enhancement in offering telecommunication services through local and international companies. By 2002 Kosovo was using mainly microwave

links [17]. However, in 2004 Kosovo enhanced its networking and telecommunication infrastructure by digitizing all switching centres, followed by the establishment of fiber optic links through Prishtina and regional centres in 2006, and in 2007 the technology was further enhanced with the creation of Next Generation Network (NGN) platform including ADSL services [18]. Currently, the network infrastructure has implemented three main fiber optic rings that interconnect six major municipalities: Prishtina- Mitrovica- Peja- Prizren- Ferizaj- Gjilan.

Current network infrastructure is built of microwave links, optic and coaxial cables [19], by some of the main Internet Service Providers such as Post and Telecommunication of Kosovo, IPKO (one of the biggest foreign investors [20]), Artmotion, Kujtesa, etc.

The network infrastructure and its layout around Kosovo is mainly dominated by the Cable/DOCSIS technology 68.95%, followed by other recent and advanced technologies, such as: xDSL 25.43%, Wireless technology 4.57%, fiber optic 0.64%, etc [21].

Major telecommunication companies operating in Kosovo provide most of the recent services, such as Televoting, IP Geocentrex, Prepaid Telephony, VOIP telephony, Internet Protocol television (IPTV), Video on Demand (VoD), PPV (Pay Per View), Video Broadcast, Music Broadcast, Music on Demand (MoD), VPN (Virtual Private Network, for example for Banks, Insurance Companies, KFOR).

The following subsections present in more detail the structure and characteristics of the networking technologies operating in Kosovo market and the Internet and smartphone usage.

5.1 Internet Service Providers

The Kosovo Telecommunication Authority has licensed almost 51 companies that provide end-to-end services and about five International Internet exchange companies [22]. More details about the number of licensed companies are presented in Table 1.

Table 1. Licensed companies by Kosovo Telecommunication Regulatory [23]

License for Telecommunication Mobile services (MNO)	**2**
License for Mobile Telecommunication services (MVNO)	**2**
License for Value Added services	**31**
License for Internet services	**49**
License for International Telecommunication services	**6**
License for Nodes and International Telecommunication equipment	**6**
License for Infrastructure of Telecommunication services	**1**
Total	**100**

As is evident, the market share is mainly dominated by four major companies. Hence, based on the latest report presented by Kosovo Association of Information and Communication Technology (STIKK) [22], the graph presented in Figure 3 presents the Internet service market share in Kosovo.
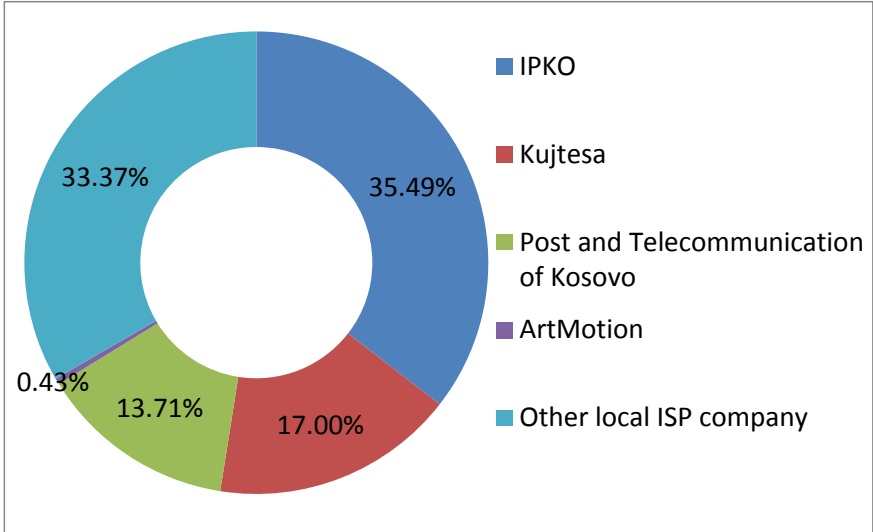
Fig.3 Internet service market share in Kosovo [22]

The local ISP providers mainly supply the rural areas, and around 44.5% of the users of rural areas are covered by local ISP companies that do not have direct access to major companies unlike the remaining 55.5% of users who do have access [22].

In Fig. 4, the estimated prices for Internet services are presented based on [22], and the bandwidth that users have at their homes. As it is shown, the highest percentage of the bandwidth that customers pay, is around 2Mbps. However, a comparable percentage is also for the available bandwidth 5Mbps. Furthermore, the figure below compares the percentages of the prices that the users pay with the available bandwidth that they get for the corresponding ISP.
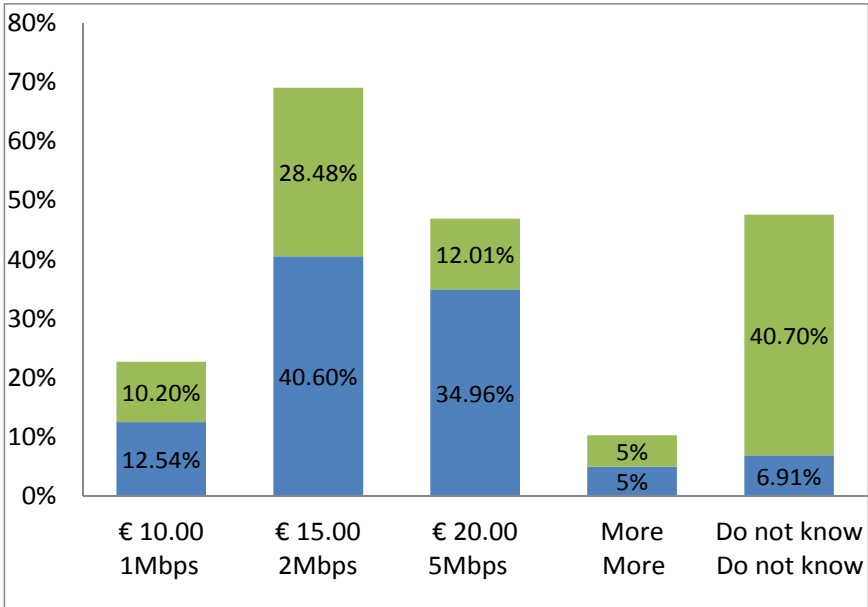
Fig.4 Price rates for provided bandwidth and Internet services [22]

As the respective figures show, the prices for the lowest bandwidth provided by operators in Kosovo which is 1 Mbps, is around 10-12 euro. This is not considered to be a high price and it is comparable to the Kosovo economic standards.

5.2   Internet Penetration

According to the latest report by the Kosovo Association of Information and Communication Technology, STIKK [22], the Internet Penetration in Kosovo is almost the same as in the developed countries, by achieving 76.6% of the users (compared to the report published by the International Telecommunication Union where the Internet penetration in developed countries was estimated to be 77%). However, the report presented in [22] concludes that if the figure is captured by households the percentage would be even higher at 84.4%. These findings are shown in Fig. 5.
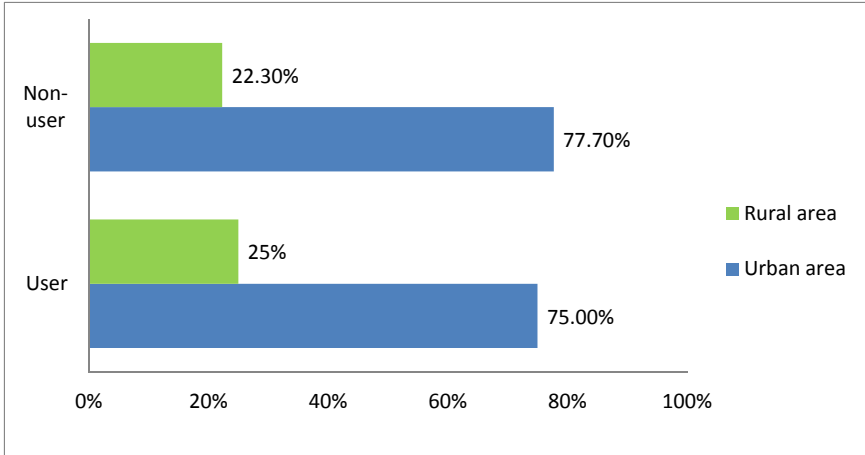
Fig.5 Internet Penetration in Kosovo based in Urban and Rural areas [22]

The report presented in [22] about the Internet Penetration in Kosovo, estimates an average of 9 wireless networks per kilometer. Furthermore, this report concludes that the Kosovo Internet users do use Internet services similarly to European citizens.

5.2.1 Smartphone Internet Usage

While considering the Internet Penetration phenomenon, it is important to understand the use of smartphones for Internet services. Based on the report presented by STIKK [22], the percentage of smartphone users who have access to Internet services are a bit higher compared to non-users, as presented in Fig. 6. The total percentage of users who use their smartphones to access the Internet is around 55.48%, where most of them, around 77.72% are of age between 10- 19, followed by users of age 20- 29.
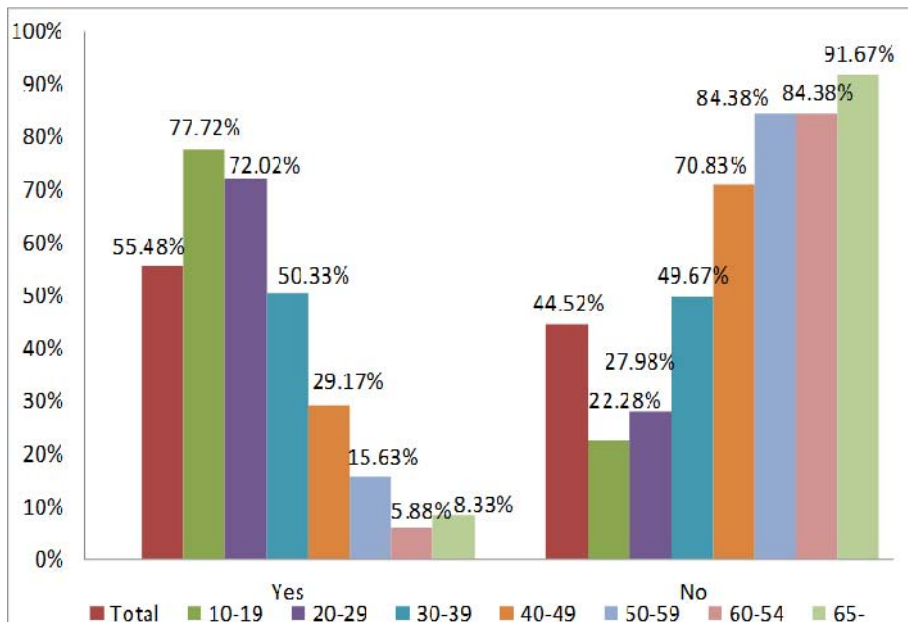
Fig.6 Percentage of Internet Access through Smartphone based on population age [22]

According to the data presented in Fig. 6, the average age of the Kosovo users is around 10-19 years old, in contrast to elderly people who do not use smartphones for Internet Services.

6   Discussion

Based on our findings and related research, the idea of designing a WSN architecture that will suite the developing region needs, is really optimistic. First of all, considering developing regions with a very limited number of healthcare institutions, and especially a limited presence of doctors and healthcare centres in rural areas, the provision of healthcare through remote communication will provide many facilities to population living there.

More precisely, taking into account the research done until now in the health system of our case study, Kosovo, the need for healthcare remote applications is present. Moreover, considering the facts drawn from this research on the high number of patients and limited number of health specialists, we believe that the use of remote healthcare providing mechanisms will help the Kosovo health system in general, and the patients in particular.

The existing network infrastructure is comparable with most of the developed regions' infrastructures. Therefore, the wireless links, the data exchange, and the security of the transmitted data can be ensured. These facts facilitate the design and the implementation of our WSN architecture.

To propose a WSN architecture, we will try to design it according to developing region needs and its economic situation. More specifically, we need to perform a cost analysis. This will lead our proposed architecture too. We will seek to include devices and elements which will form a suitable and affordable architecture for a developing country.

7 Conclusions

There are many advantages that WSN applications bring to a developing country. Therefore, in seeking to propose a WSN architecture and develop the corresponding application, we have analyzed the current trends, infrastructure and need for such an application. More specifically, considering our case study of Kosovo, we made an estimation of the patients' need for continuous monitoring, and the telecommunication infrastructure in order to implement the WSN architecture. The research results are promising, in the context that Kosovo health system requires improvement and it needs to lower the number of hospitalized people. On the other hand, the current network infrastructure is comparable with the networks of developed countries.

References

1. Vo, A., Brook, G.B., Farr, R., Raimer, B.: Benefits of Telemedicine in Remote Communities & Use of Mobile and Wireless Platforms in Healthcare. UMTB Telemedicine and Centre for Telehealth Research and Policy. Paper sponsored by: Internet Innovation Alliance
2. Zennaro, M., Pehrson, B., Bagula, A.: Wireless Sensor Networks: a great opportunity for researchers in Developing Countries. In: 2nd IFIP Intl. Symposium on Wireless Communications and Information Technology in Developing Countries. South Africa (2008)
3. Balakrishna, D., Sujeethnanda, M., Rama Murthy, G.: Mobile Wireless Sensor Networks: Healthcare in Hospitals. In: Fifth International Conference on eHealth, Telemedicine, and Social Medicine. Report No: IIIT/TR/2013/-1. INDIA (February 2013)
4. Aminian, M., Naji, H.R.: A Hospital Healthcare Monitoring System Using Wireless Sensor Networks. In: J Health Med Inform 4: 121. Volume 4, Issue 2, 1000121(2013)
5. AirStrip ONE OB: Patient Safety and Physician satisfaction, AirStrip White Paper Series, http://www.airstrip.com/sites/default/files/AirStrip_PS_AirStrip_ONE_OB_v2.pdf
6. Lee, D.H., Rabbi, A., Choi, J., Fazel-Rezai, R.: Development of a Mobile Phone Based e-Health Monitoring Application. International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 3, No. 3 (2012)
7. Qualcomm: Qualcomm and Life Care Networks Launch 3G Mobile Health Project to Help Patients with Cardiovascular Diseases,

https://www.qualcomm.com/news/releases/2011/09/07/qualcomm-and-life-care-networks-launch-3g-mobile-health-project-help , (September 7, 2011),

8. West, D.M.: Improving Health Care through Mobile Medical Devices and Sensors. Article by the, Centre for Technology Innovation at Brookings (October 2013)

9. Kosovo Statistical Agency, http://ask.rks-gov.net/vitalstatistics (2012)

10. Kosovo Statistical Agency: Kosovo Population 2012, https://ask.rks-gov.net/publikimet/doc_download/993-vleresim-popullsia-e-kosoves-2012 (April 2013)

11. Kosovo Statistical Agency.: Kosovo population projection 2011- 2061. Technical report. Prishtina (December 2013)

12. Hostl, J..: Health in Kosovo- A general view (Shëndetësia në Kosovë- Një pasqyrë e përgjithshme). Part of the evaluation report committed under the Ministry of Foreign of Luxembourg, version 15.05.07

13. IPOL, Balkan Policy Institute. : Research about the health system in Kosovo (Hulumtim në sistemin shëndetësor të Kosovës). Research article financed by: UNDP,United Nations Development Programme - UNDP project "Parliamentary Development for Social Policies", Prishtina (September 2012)

14. Public Pulse.: Satisfaction with health services and perception about the presence of corruption (Kënaqshmëria me shërbimet shëndetësore dhe përceptimet mbi praninë e korrupsionit). Research financed by USAID and UNDP (November 2013)

15. Buwa, D., Vouri, H.:Rebuilding a health care system: war, reconstruction and health care reforms in Kosovo. European Journal of Public Health, 17: 226- 230 (2006)

16. Percival, V., Sondorp: A case study of health sector reform in Kosovo. Conference of Health, 4:7 (2010)

17. Caka, N., Cakaj, Sh.: The development trends of the transmission systems of Kosovo telecommunications network. Sustainable technical-technological development and Environment, pp. 261-268, Prishtina (2002)

18. Post and Telecommunication of Kosovo (PTK) J.S.C., http://www.developingmarkets.com/sites/default/files/session-2_2-dr-shyqri-haxha.pdf

19. Analysis of ICT industry in Kosovo. Kosovo Economic development through quality and networking. Research report financed by European Union Regional Economic Development programme and European Union Office in Kosovo (2011)

20. Telegeography, http://www.telegeography.com/products/commsupdate/articles/2006/05/30/telekom-slovenije-buys-75-of-kosovan-isp/

21. Telecommunication Regulatory Authority: Annual Report 2011, http://www.art-ks.org/repository/docs/Raporti%20Vjetor%202011%20i%20Autoritetit%20Rregullator%20te%20Telekomunikacionit1.pdf (2011)

22. The Kosovo Association of Information and Communication Technology (STIKK): Internet Penetration and Usage in Kosovo. Annual report, (August 2013)

23. Regulatory Authority of Electronic and Postal Communications, http://www.art-ks.org/repository/docs/Regjistri%20i%20Ndermarrjeve%20te%20Licencuara.pdf

Social Network Application for Health Care Delivery

Marika Apostolova Trpkovska[1], Betim Cico[2], Ivan Chorbev[3]

[1,2] Faculty of Contemporary Sciences and Technologies

South East European University, Tetovo, Macedonia

[3] Faculty of Computer Science and Engineering

Ss. Cyril and Methodius University, Skopje, Macedonia

[1] m.apostolova@seeu.edu.mk, [2] b.cico@seeu.edu.mk, [3] ivanchorbev@gmail.com

Abstract. Social media use in health remains an issue with small evidence support. Conventional media promotions are used broadly in public health for a large variety of objectives and have shown positive outcomes in terms of public awareness. Social media networking's' are becoming powerful addition to the health issues. There is a huge interest in using social media as a tool for public health matters. The aim of this paper is to create a social network that will be used by young mothers not only to exchange experiences through different social services but also to get necessary information about disease predictions and preventions, immunization tracking and other matters. The social network would exist as the first point where mothers could gather health oriented information in their mother tongue. This social network would be especially designed for mothers that live in rural areas and cannot get instant medical help, information and advice.

Keywords: social network, health care delivery, children health, health communications.

1   Introduction

Social media continue to grow globally offering individuals fresh and significant ways to engage with the public and issues that matter to them. According to Nielsen and NM Incite's latest Social Media Report, consumers continue to spend more time on social networks than on any other category of sites [1].

Social media are the process of people using online tools and platforms to share content and information through conversation and communication. They are drastically changing the way of communication and we should not take too lightly their ability to work for organizations. Globally, a variety of organizations, especially the health segment, are already using social media as an important tool to connect consumers and providers as well as to inform product development [2], [3].

The healthcare organizations have started to turn to social media as devices for connecting with the public audience, as more individuals turn to the Internet for health related issues. As a rising topic, very few existing studies have analyzed the effectiveness of social media in this perspective. More studies are needed before social media's role in public health can be adequately defined although preliminary researches have demonstrated significant results related to social media applications that have potential for engaging precise target audiences. This is due to the difficulties in social media applications evaluation and the lack of publishing such reports remains to be the main barrier for gathering proof of their value. Regardless of this, public health organizations are already using social media applications, and some are starting to launch them. In spite of the fact that the published evaluations of these initiatives are basic, they are very useful. In these evaluations, the social media are very practical alternatives that are fairly easy to adopt, but require adequate and proper human resources to maintain, precise business plan and guidelines for their use in a public health context [4].

For this reason we would like to propose social media to be included in Macedonian health system to enhance the health care delivery and help. We propose this network to be especially designed for young mothers, where they could stay in touch, share parenting advice, review products, create groups, discuss topics and activate different health oriented services like disease prediction tool, vaccination calendars and others.

This social network would especially help mothers that live in rural areas, not just to share experiences with other mothers through diverse available features and services, but also to get some instant opinions of what might be wrong with their children's health, based on some visible symptoms, as well as to be reminded about their kids' vaccinations. The prediction tool would be used as a first point where mothers, based on few present symptoms, can be informed about the possible problem that the child has. After that, the consultation with real medical specialist must be done in order to determine the exact disease and to continue with the treatment.

Fig. 1. eMama social network site

As we can see from the Fig. 1, we called this social network **eMama** and it is already available on the following link *http://www.emama.mk/*. Because of the testing phase, all the babies' pictures in *www.eMama.mk* are taken from the online public free domain pictures repository.

2   Social Media Practices

Before starting to realize our idea we did few preliminary Internet researches to see the actual impact of different kinds of media on human life. What we get as a result is that media are playing a major role in informing various aspects of peoples' lives. It was found that health organizations usually use print, radio and television media to distribute essential health information to the public audience [5]. Over the past decades the electronic media have become the leader in communicating. Knowing this, public health communication has evolved to reflect to it and their services are becoming increasingly digitalized. In the stage of "Web 2.0" and wake of the "Web 2.0" phenomenon, public health communication strategies are also shifting to go with the increasingly powerful and rapidly evolving social media revolution [6].

Social media appearance has also deeply changed the way individuals interact in the society, engaging people both online and otherwise. Social networking sites like Facebook, LinkedIn, Twitter, PatientsLikeMe and others allow individuals to communicate with large groups of friends, contacts and colleagues [7]. Health information gathering is no immune to this trend. Social media provide an outlet for the publication of health information to consumers, while allowing consumers to give their thoughts and contribute to recommendation that was usually only issued by health departments [8], [9].

Social media respond to the needs of modern health consumers, who are immensely engaged and aspire to be more involved in their health decision than at any point previously, while providing the means that was previously unbelievable [10]. To some extent, various private and public organizations are paying attention to social media use because of the huge potential they provide in reaching target members through rapidly disseminated messages [11]. For instance, pharmaceutical companies are using social media to promote their drug brands [12], other to promote businesses brands [26], to predict future events [13], to predict and track diseases [14] etc.

Having these advantages, a variety of healthcare organizations and units are also adopting social media in their everyday practices in a variety of ways, for outreaching the community, for educating the patients, for marketing the products and for communications [15], [16]. The Mayo clinic center, has developed a patient-centered social networking site that allows patients to interact with healthcare professionals (similar to Facebook), but with unique and private interface. The consumers had a huge interest and almost 1000 users joined their network within the first 7 days. They think that they are strengthening their health care needs using these social media tools [17].

Another report particularly made by the Region of Peel-Canada Public Health department, intends to review the existing and potential uses of social media on the Public Health sector. The report highlights the move towards participatory methods for health consumers from the traditional protectors of health related information. Despite the fact that huge numbers of organizations are currently using social media, there is a deficit of experimental or quasi-experimental evidence on its utility. The existing ones are to observational and are just analyzing the current health related communications through social media [18].

In addition to this, public health is also taking into consideration the accomplishment of social media by Infoveillance, which is a syndromic surveillance that utilizes online contents. Health organizations can as well use social media tools for this purposes. They can monitor the frequency of searches related to a particular disease, requesting the consumers to report appeared symptoms. Afterwards they can map occurrences with tools and data gathered from the social networking sites to provide critical information [19].

Next, the health information quality available through social media is also a crucial element. The public health organizations mentioned the capacity to provide trustworthy information to consumers available at any time, place and format. It is also essential to reach target audiences [20].

The main obstacle in adapting social media technologies lie in the perceived resistance of government organizations to change procedures and policies needed to begin and effectively sustain a social media presence [21]. Although the existing literature supports the possibility of public health issues being delivered through social media, more information is needed to verify the real implementation costs [22].

Being aware of this, our idea to incorporate social media for helping mothers get any kind of information related to health, would be a step forward for Macedonian health system where our network could be a part of their system. This will help young mothers to be informed about any issues related to their children's health. Through our network essential health information to the public audience can be distributed. Now the mothers can give their thoughts and contribute to recommendation that was usually only issued by health departments in our country. Also mothers would be more involved in their children's health decision than at any point previously done. What we can say for sure is that eMama.mk would strengthen mothers' health care needs in any aspect.

3 Social Media in Macedonian Health Oriented Communications

Every new system designed needs to be dedicated to someone or something, especially when the health is being concerned. So, what we propose can be in general but we propose the case study to be Macedonia and more in general the Macedonian young mother.

To address the necessity of having social media incorporated in human everyday living a qualitative interview and anonymous phone calls were used to explore the

mothers' perceptions about the impact of social media in health. They were conducted among young mothers from Tetovo and several doctors-pediatricians from state and private hospitals. The survey provided a sufficient number of responses. Actually, the survey records representative response from 150 young mothers from Tetovo and 6 pediatricians (3 from Tetovo State Clinic and 3 from private hospital in Skopje - Remedika). The actual responses, as supposed, gave a high level of supporting the idea of incorporating the social media in health care by both sides. On the question: *Do you think that there is a necessity in Macedonia for creation of a social network for finding any kind of children health oriented information in native language?*; we did not receive any negative response as we can see from the Fig. 2.
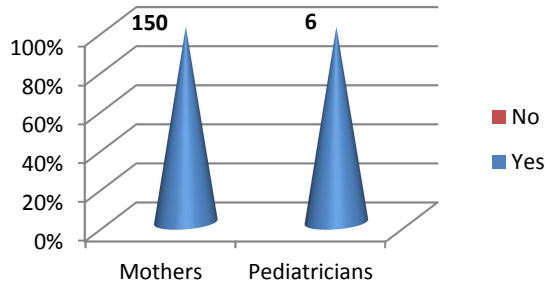
Fig.2. Data results from the interview question

Having all the advantages of the use of social media and their attractiveness among the population, we are proposing this social element to be part of the design of health oriented communications within a framework of Health system development undergoing in Macedonia. Under health oriented communications the attention can be put on promoting campaign messages and organizations' activities (like disease preventions details, immunization info, wellness, drugs deliverance, health professional education and collaboration, marketing etc.).
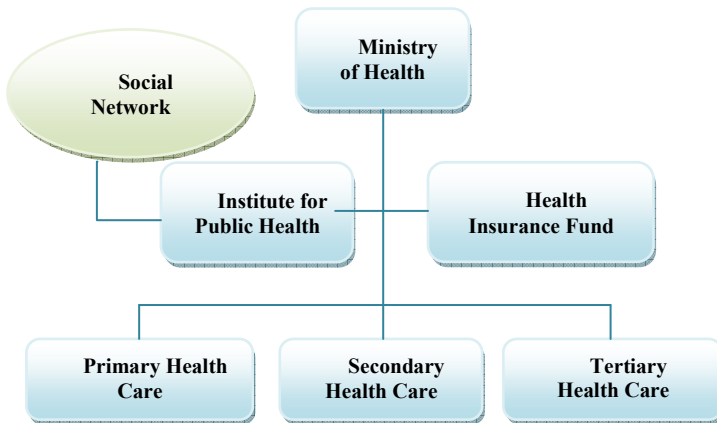
Fig. 3. The actual health system in Macedonia

Therefore, the primary idea is to incorporate social media presence in the scope of Macedonian health system framework. In the Fig. 3 we are presenting the actual organization structure of the health system in Macedonia [23].

Analyzing all the main actors of the Macedonian health system, the social media can be taken under the Institute for Public Health of the Republic of Macedonia [24] leadership. IPHs' aim is to implement the national strategy in presenting the components of health and medical activities for improving the health of the inhabitants of Macedonia. A special office could be assigned as main interface between our social network and e-health system in Macedonia, Fig. 3.

Actually our idea is that the social network eMama.mk exists as a first point where mothers could gather needed health oriented information. What makes our proposal different from the others is the simplicity and cover of varied segments of the motherhood life in one online place. We strive to make mothers' life easier when there is a need of getting online health information.

Actually, the idea is that social network is used by young mothers not only to exchange experiences through different social services but also to get necessary information about disease predictions and preventions, immunization issues, wellness, drugs deliverance, review products, to create groups, to discuss topics, to have real time chatting and other issues in the mothers' native language. The other thing is engaging precise target audiences and functional point, mothers and health. Here the mothers could be informed, could communicate and collaborate about different health problems which are the most valuable segment of the human life, especially when they are connected with children.

The social network ties would be limited to people that are very much alike. For example, a mother whose child suffers from a particular genetic disease will try to get help from someone else that has experienced the same problem, to get the information needed, to stay in touch and share advice.

This network place will also offer different health oriented services like vaccination calendar (which is in process of designing but still available on the following link: *http://vakcina.gojadizajn.com.mk/)* and it will have a specialized Children Disease Prediction Engine implemented, based on our MS2TP model [7]. The prediction engine is based on our proposed model [25]. Now we are working on mapping the social network questionnaire data with the specially designed ontology which is the main part of the proposed model.

Namely, the proposed model goes through several phases. In the first phase, after the user fulfills the available survey data about the disease in the social network, they would be inserted in the main database. The survey would consist of no more than 10 questions related to the disease that can be easily fulfilled by the user. The flow of activities of prediction survey process is shown in Fig.4. Here, the user fills in some important disease data and after the user starts to enter the appeared symptoms. After the user finishes entering data, the prediction process starts by clicking on Predict button where the second phase of the model begins. We will use relational databases to define a bridging mechanism between relational database and OWL ontology. The ontology and the database have been developed separately. In the next phase the mapping process will start between database/ontology structures using D2RQ Server. In the last phase we are going to use semantic reasoner for checking the reliability of the relationship among the classes and their properties, in order to validate the SWRL

rules to detect disease from the given symptoms. So, after survey data are entered, they are mapped with our ontology and the system requests the query processor. The query processor checks with SWRL rules for relations between the diseases. It returns the diseases associated with symptoms and other data entered by the user and after that displays the output to the user [25].
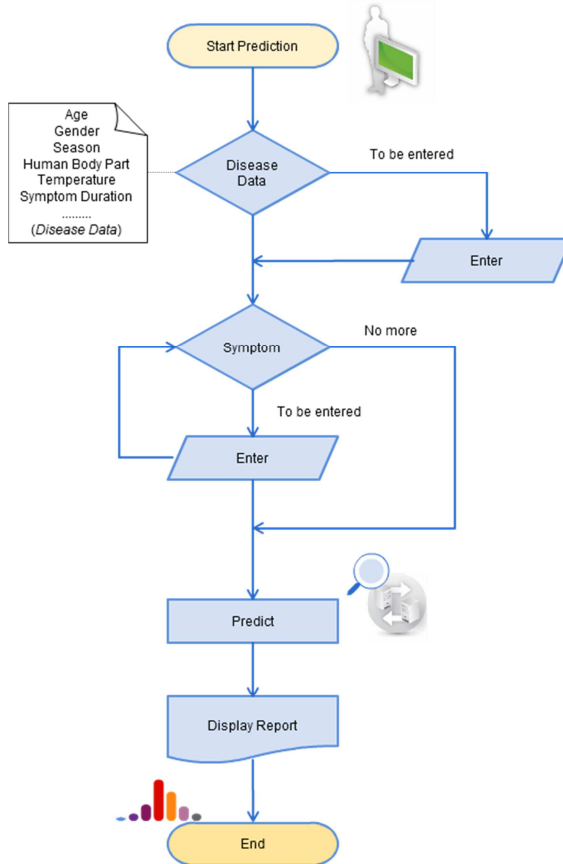
Fig. 4. Flow of activities of the prediction process

4 Conclusion and future Works

Social media are here. Various healthcare organizations are using them as significant tools to connect consumers and providers. The social media phenomenon overtakes the information revolution and drastically changes the way we communicate and achieve healthcare goals.

Social media tools, as proposed in this paper, can be utilized as part of an integrated health communications program for a wide variety of objectives and positive outcomes, offering an important set of tools to open communication between providers and consumers, often by going to where consumers are. As soon as

organizations develop an active social media presence, the less distance they will have to make up later. Now is the time when Macedonia could be a leader in providing technology-enabled solutions and services for the consumers.

At this point, we are testing and implementing eMama.mk to determine its endorsement. Various evaluations would be performed to show the successful execution of the proposal.

Health is a very sensitive issue and a social network about it requires very thoughtful considerations on how to tackle issues like privacy, trust, confidentiality and security. These issues would be taken into consideration in our future research.

In addition we would also like to see our social network solution to work on mobile and cloud.

References

1. Nielsen.: Social media report 2012 - social media comes of age, http://www.nielsen.com/us/en/newswire/2012/social-media-report-2012-social-media-comes-of-age.html (2012)
2. Pharma Voice.: Healthcare and social media: A winning formula, http://www.mckinsey.com/~/media/McKinsey/dotcom/client_service/Pharma%20 and%20Medical%20Products/PMP%20NEW/PDFs/PharmaVOICE0112-Showcase-McKinsey.ashx (2012)
3. Pharma Voice.: The mHealth Revolution, http://viewer.zmags.com/ publication/6b41916c#/6b41916c/26 (2011)
4. MIEH.: McMaster Institute Environment and Health Report, http://www.mcmaster.ca/mieh/documents/publications/Social%20Media%20Repo rt.pdf (2011)
5. Unite for Sight.: The Role of Media in Health Promotion, http://www.uniteforsight.org/health-communication-course/module5
6. Schein, R., Wilson, K., Keelan, J.: Literature Review On Effectiveness of the Use of Social Media a Report for Peel Public Health, http://www.peelregion.ca/health/resources/pdf/socialmedia.pdf
7. Apostolova, T.M., Cico, B., Chorbev I.: Using Social Media to Predict Children Disease Occurrence. Proceedings of 7th IADIS 2014, pp. 209 -216. Madrid, Spain (2014)
8. Avery, E., Lariscy, R., Amador, E., Ickowitz, T., Primm, C., Taylor, A.: Diffusion of social media among public relations practitioners in health departments across various community population sizes. Journal of Public Relations Research. 22(3):336-58 (2010)
9. AcademyHealth.: Improving Quality Health Care: The Role of Consumer Engagement. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rj a&ved=0CCUQFjAA&url=http%3A%2F%2Fwww.academyhealth.org%2Ffiles% 2Fissues%2FConsumerEngagement.pdf&ei=O0nqUq3CMKON4wTm0oDwCg& usg=AFQjCNFt6YNn5Zu-1XPZWOIplufkQGRQMQ&sig2=yNfcogIg0Zt8L8G8RNdh8Q&bvm=bv.604445 64,d.bGE

10. Schein, R., Wilson, K., Keelan, J.: Literature review on the effectiveness of the use of social media: A report for peel public health (2010)
11. Edward, S., Nichols, J.: Engaging a new community: Going grassroots with facebook. Transfusion. Conference: AABB Annual Meeting and CTTXPO Baltimore, MD United States. Conference Publication: 50:106A (2010)
12. Barton, A.: Big pharma wants to 'friend' you. Globe and Mail (2011)
13. Efthimios, B., Dimitris, A., Gregoris, M.: Using Social Media to Predict Future Events with Agent-Based Markets. IEEE Intelligent Systems, vol. 25, no. 6, pp. 50-58 (2010)
14. Charles, W., Schmidt, C.W., Schmidt, M.S.: Trending Now: Using Social Media to Predict and Track Disease Outbreaks. Discover Magazine, Science, and Nature Medicine,
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3261963/
15. NCBI.: Networking Health: Prescriptions for the Internet,
http://www.ncbi.nlm.nih.gov/books/NBK44715/
16. Mayo Clinic Center.: Social Media Health Network, A Service of the Mayo Clinic Center for Social Media, http://ebennett.org/hsnl/
17. Mayo Clinic.: Mayo clinic launches social network to connect global mayo clinic community, http://socialmedia.mayoclinic.org/2011/07/05/mayo-clinic-launches-social-network-to-connect-global-mayo-clinic-community/
18. Schein, R., Wilson, K., Keelan, J.: Literature review on the effectiveness of the use of social media: A report for peel public health (2010)
19. Shweta, A., Mayanka, G., Swati, K.: Social Media And Effective Health Communication. International Journal of Social Science & Interdisciplinary Research ISSN 2277 3630. IJSSIR, Vol. 2 (2013)
http://indianresearchjournals.com/pdf/IJSSIR/2013/August/5.pdf
20. Miriam, J.M.: Making Sense of Credibility on the Web: Models for Evaluating Online Information and Recommendations for Future Research. WILEY InterScience (2007)
http://www.ischools.org/conference08/pc/WC17_iconf08.pdf
21. NCBI.: Networking Health: Prescriptions for the Internet.
http://www.ncbi.nlm.nih.gov/books/NBK44715/
22. Bennett, G.G., Glasgow, R.E.: The delivery of public health interventions via the internet: Actualizing their potential. Annu Rev Public Health. 30:273-92 (2009)
23. CRPM, Centar za Istrazuvanje I Kreiranje Politiki,
http://www.crpm.org.mk/wp-content/uploads/2012/03/KnigahealthMAK1.pdf
24. Institute for Public Health of the RM,
http://www.iph.mk/index.php?lang=en
25. Apostolova, T.M., Betim, C., Ivan, C.: Application of Social media in eHealth. MECO'2014. ISBN 978-1-4799-4826-0, pp. 314-317 (2014)
26. Using Social Media to Promote Your Brand, Grow your Network & Generate Revenue. http://sparkandhustle.com/takeaway-tips/using-social-media-to-promote-your-brand-grow-your-network-generate-revenue

Towards Mobile Semantic Decision Tables for Personalised Semantics

Ioannis Arampatzis

Web & Information Systems Engineering Lab
Vrije Universiteit Brussel
Pleinlaan 2, 1050 Brussels, Belgium

South-East European Research Centre
City College - International Faculty of the University of Sheffield
24 Proxenou Koromila Street, 54622 Thessaloniki, Greece
iarampatzis@seerc.org

Abstract. With the proliferation of intelligent environments, devices increasingly operate in a context-aware manner. What existing research efforts lack however, is the use of personalised semantics. In this work we adopt an ontology-oriented approach as a modelling technique for context-aware applications. We propose an architecture for the management of smart environments together with the use of Semantic Decision Tables (SDT). We furthermore implement an important component for this architecture, which is the mobile SDT editor. The latter works as a rule configurator and facilitates users to create and use personalised semantics.

Keywords: Ontology Engineering, Semantic Decision Tables, Context-Awareness, Intelligent Environments, Smart Home Management, Android

1 Introduction

Automation of a smart environment is the use of information technology to control the available smart objects that comprise it [1]. The latter are autonomous physical or digital objects augmented with sensing, processing, and network capabilities [2]. This enables remote control through the use, for example, of mobile devices. During the last decade several models of intelligent environments have been proposed [3]. The work in [4] suggests that in order to operate in a dynamic environment, e.g. a smart home, devices must operate in a context-aware manner. The term "context" refers to *"any information that can be used to characterise the situation of entities that are considered relevant to the interaction between a user and an application"* [1].

According to [5], there are two prerequisites for building context-aware applications: the context model and the middleware-level support. The latter refers to the data that are obtained from the physical layer (sensors) and the database.

Regarding the context model, there are three approaches to modelling techniques proposed in [6]: the application-oriented, the model-oriented, and the ontology-oriented. As opposed to the latter approach, the application oriented and model-oriented approaches do not support knowledge sharing or context reasoning. The use of ontology languages such as RDF(S) and OWL enables the specification of concepts and their interrelations, as well as reasoning about their properties [6]. More specifically, reasoning supports the creation and the management of user-specified rules in order to control the smart objects of an intelligent environment.

In this work we propose an architecture for the management of business rules and, in particular, decision rules through the use of Semantic Decision Tables (SDT) [7]. An SDT is annotated and formalised through the use of ontologies, thereby overcoming the lack of semantic support that exists in decision tables. Unlike other decision support tools, SDTs are business-oriented, and thus suitable for being used by non-technical people.

In this paper we use SDTs in a use case scenario for the management of smart homes. Through the use of the proposed mobile SDT editor, users are greatly facilitated in formulating decision rules, as they are able to express them in natural language. The SDT editor translates the created SDTs in XML files of a specific XML Schema, which contain the information regarding the decision rules, as well as the annotated conditions and actions that apply to the relevant ontology. This way, users take advantage of the flexibility, shareability, and formality of SDTs.

This paper is structured as follows. Section 2 motivates the presented work and outlines relevant research. Section 3 provides the background knowledge regarding SDTs. Section 4 presents the architecture of our system and the ontology. Section 5 outlines the implementation aspects. Finally, Section 6 presents results and future work.

2 Related Work & Motivation

A multitude of architectures that utilise semantic technologies for supporting context-aware applications in smart environments have been proposed [8]. These lack, to the best of our knowledge, components to support personalised semantics, a feature offered through the SDT editor by the architecture proposed in this work.

SESAME-S [9] is the most relevant project to our work. It adopts an ontology-based modelling approach whereby rules are translated into semantic policies. The proposed architecture consists of a central Universal Control Box that brings together an ontology and a rule infrastructure for context and rule-based reasoning. The main purpose of this project is to reduce the users' energy consumption. End users are being offered a collection of pre-defined rules by power users, in order to avoid incorrect decisions. Power users are domain experts responsible for devising the device models. The companies (vendors) that provide the sensors and the actuators could be considered power users. In contrast to SESAME-

S, our implementation allows the definition of decision rules by all users (even by non-experts), as well as the linking of these decision rules to appropriate ontologically-expressed properties to support interoperability and shareability of the action plans.

The proposed implementation provides a number of advantages. Firstly, decision tables are easy to learn and to construct. This means that we offer a simple user interface, facilitating the creation of decision rules. Secondly, through the use of natural language and the intuitive layout of SDTs, we offer users the ability to easily manage (modify, read) their decision rules. Thirdly, stored decision rules dispense with the requirement of users managing their smart objects in real time. This is because once users create their management plans via the SDT editor, the smart objects will behave accordingly, every time the pertinent SDTs are uploaded on an SDT-enabled server. Last but not least, users will be able to share decision rules via an SDT repository.

3   Semantic Decision Tables

A proliferation of decision support technologies exists, including decision tables, decision trees, Bayesian nets, balanced score cards, decision models and SDTs. An SDT is based on the DOGMA approach [10]; it embodies a decision table comprising three parts: conditions, actions, and rules. Conditions comprise a condition stub and a condition entry. Similar to conditions, actions comprise an action stub and an action entry. Rules are a combination of conditions and actions; an example of a decision table is provided in Table 1.

Table 1: Example of a decision table

Condition	1	2	3	4
Temperature is high (>25)	Yes	No	Yes	No
Low relative humidity (<40)	Yes	Yes	No	No
Action	1	2	3	4
Water the plant	*			

An SDT also includes a set of lexons, and a set of commitments. Lexons are expressed in terms of 5-tuples, for example $\langle c, Actuator, performs, perfomed by, Action \rangle$. This lexon represents a fact that is in the context identified by c where *Actuator* and *Action* are defined. *Performs* and *performed by* are the two roles that the previous two concepts can undertake. From the example above, the lexon instances that can be derived from the lexon base are:

- $\langle c, Sensor, with, of, SensorType \rangle$
- $\langle c, Sensor, with, of, ObservationValue \rangle$
- $\langle c, Actuator, performs, perfomed by, Action \rangle$
- $\langle c, Action, with, of, ActionName \rangle$

Commitments are statements that make use of lexons and are used to specify and to annotate constraints. Commitments are beyond the scope of this paper; more information can be found at [11].

4 Architecture

Figure 1 illustrates the proposed architecture. The D2RQ[1] platform provides access to the data of the relational database via virtual RDF graphs. The data are mapped to the pertinent ontology according to the mapping file of the platform. The main advantage of using this architecture is that we can query the relational database using a SPARQL endpoint, and also that we can access the content as Linked Data [12]. The server also acts as an SDT repository.

To access the data from the SDT editor, we use the Androjena[2] API, as it is the only framework that offers a query engine for mobile applications with the use of ARQoid[3]. The SDT simulator is a tool for users to check that the SDTs they have generated behave according to their expectations. In this sense, it is a tool that simulates smart environments.

4.1 Ontology

A conceptual schema of our ontology has been built using the NORMA[4] modelling tool. NORMA is a plug-in for the Microsoft Visual Studio, and is a tool for conceptual modelling based on the Object-Role Modelling (ORM) approach [13]. Through NORMA the relational database is generated, and is used to generate the D2RQ mapping file. To create our ontology we use GOSPL [14], which is a hybrid ontology engineering method and tool. To this end, GOSPL provides a link[5] to access the OWL implementation of our ontology.

5 Implementation

In this section we analyse the two components of our architecture, the mobile SDT editor and the SDT simulator.

5.1 Mobile SDT Editor

The Android application, referred to the SDT editor in this work, is the application for the management of the decision rules. Using this application users are able to create and edit the required conditions and actions by using natural language through the use of their mobile device.

---

[1] http://www.d2rq.org
[2] http://code.google.com/p/androjena
[3] http://code.google.com/p/androjena/wiki/ARQoid
[4] http://http://www.ormfoundation.org/
[5] http://starpc18.vub.ac.be:8080/gospl/ontology/62
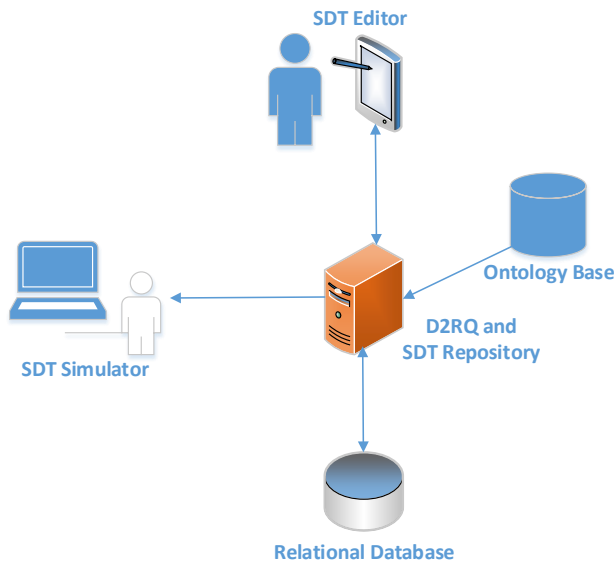
Fig. 1: System Architecture

The SDT editor consists of four activities: the general management activity, the conditions management activity, the actions management activity, and the SDT creation activity.

General Management Activity. It is the activity where the user specifies the URL to the SPARQL endpoint. For the purposes of this paper, we suppose that the initial setup of the smart objects is already inserted to the relational database. This activity offers the ability to the user to download or to update relevant data from the server. Last but not least, the user is able to manage the SDT-generated files from those that are already submitted and stored locally on the client's side.

Conditions Management Activity. It lists all the conditions that are downloaded from the server (Fig. 2). This includes the condition stub name, the condition entry, and the property that the condition refers to. This way, when the user specifies the conditions in the SDT, only boolean values have to be inserted, as the propositions are pre-defined via the conditions management activity, forming what is termed second-order SDTs [15].

Actions Management Activity. It lists all the actions that the user can perform in his smart environment. The actions management activity is responsible for mapping natural expressions into action identifiers. For example, in Table 1 the action "Water the plant" is mapped with the actuator that is responsible for enabling the automatic watering system.

Smart Conditions Management

Property	Condition Stub	Range
Light	not enough light	<=80 %
Humidity	too dry	<=40 %
Temperature	too hot	>=20 °C
Temperature	too cold	<20 °C
Light	too much light	>80 %

Add	Delete

Fig. 2: Conditions Management Activity

SDT Creation Activity. It is responsible for generating the SDT XML file. Firstly, the user selects the desired conditions and actions from those that are specified in the previous activities (Fig. 3). Following the selection, an SDT is automatically generated and the user can specify the desired rules. The specification of the rules is supported by using checkboxes for the action entries (Fig. 4). SDT updates through adding and removing conditions and actions is facilitated. When the decision rules are inserted, users submit the SDT and an XML file is automatically generated and stored locally. This XML file includes the information from the SDT together with the annotations to the relevant ontology.

An important feature of the SDT editor is that users can share the created SDTs. The SDT XML files can be uploaded to the server, in the SDT repository, and accessed by the users.

5.2 SDT Simulator

The SDT simulator is a java-based application. The purpose of this application is to simulate an intelligent environment and to facilitate users in checking whether the generated SDTs behave according to their expectations. Firstly, a file is selected from the offered list. The conditions that pertain to this file are then listed and users can select the ones they wish to apply. According to the selected conditions, a list with the actions that are triggered from the applied decision

Please select the Conditions and the Actions

	Name	Property	Stub
☐	not enough light	Light	Condition
☑	too dry	Humidity	Condition
☑	too hot	Temperature	Condition
☐	too cold	Temperature	Condition
☐	too much light	Light	Condition
☐	Drag flower to shadow	-	Action
☐	Open window	-	Action
☐	Open the curtain	-	Action
☐	Close window	-	Action
☑	Water the plant	-	Action
☐	Drag flower to window	-	Action
☐	Close the curtain	-	Action

Select All	Reset	Continue

Fig. 3: Conditions and Actions selection

rules is displayed. This way users can check the decision plans and how the applied decision rules affect the management of the smart environment.

6 Conclusions & Future Work

The proposed architecture provides users with components that support natural language for the automation of smart environments. Furthermore, through the use of the ontology, shareability and interoperability are facilitated. The implemented components offer an easy-to-learn and easy-to-use interface, features that, we believe, will promote the widespread use of our architecture. In the future we intend to devise a methodology for the collaborative construction of SDTs, on the basis of consensus inside stakeholder communities. More specifically, as SDTs and GOSPL are both based on the DOGMA approach, a method for the collaborative construction of SDTs by the use of GOSPL will be investigated. This way, GOSPL communities' stakeholders will be able to collaboratively construct decision rules.

Regarding the SDT editor, it offers the creation of an SDT up to four conditions due to size limitations of the screen of the Android devices (in our tests we have used a tablet device of 10 inches).

Last but not least, our ontology will not stick to that implementation and will be furthermore enriched with existing ontologies, e.g. the Semantic Sensor Network ontology (SSN) [16].

CONDITION				
too hot (>=20)	Yes	No	Yes	No
too dry (<=40)	Yes	Yes	No	No

ACTION				
Water the plant	✔	☐	☐	☐

Fig. 4: Creation of SDT

References

1. A. K. Dey, G. D. Abowd, and D. Salber, "A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications," pp. 97–166, 2001.
2. G. Kortuem, F. Kawsar, D. Fitton, and V. Sundramoorthy, "Smart Objects as Building Blocks for the Internet of Things," *Internet Computing, IEEE*, vol. 14, no. 1, pp. 44–51, 2010.
3. C. Bolchini, C. A. Curino, E. Quintarelli, F. A. Schreiber, and L. Tanca, "A Data-oriented Survey of Context Models," *ACM Sigmod Record*, vol. 36, no. 4, pp. 19–26, 2007.
4. B. Schilit, N. Adams, and R. Want, "Context-Aware Computing Applications," in *Mobile Computing Systems and Applications, 1994. WMCSA 1994. First Workshop on*, pp. 85–90, IEEE, 1994.
5. V. Sugumaran and J. A. Gulla, *Applied Semantic Web Technologies*. CRC Press, 2011.
6. T. Gu, X. H. Wang, H. K. Pung, and D. Q. Zhang, "An Ontology-based Context Model in Intelligent Environments," in *Proceedings of communication networks and distributed systems modeling and simulation conference*, vol. 2004, pp. 270–275, 2004.
7. Y. Tang and R. Meersman, "On Constructing Semantic Decision Tables," in *Database and Expert Systems Applications*, pp. 34–44, Springer, 2007.
8. A. Bikakis, T. Patkos, G. Antoniou, and D. Plexousakis, "A Survey of Semantic-based Approaches for Context Reasoning in Ambient Intelligence," in *Constructing ambient intelligence*, pp. 14–23, Springer, 2008.
9. A. Fensel, S. Tomic, V. Kumar, M. Stefanovic, S. V. Aleshin, and D. O. Novikov, "SESAME-S: Semantic Smart Home System for Energy Efficiency," pp. 46–57, 2013.
10. M. Jarrar and R. Meersman, *Ontology Engineering - The DOGMA Approach*, vol. 4891, ch. 3, pp. 7–34. Berlin, Heidelberg, Germany: Springer-Verlag, 2008.
11. Y. Tang and R. Meersman, "Towards Building Semantic Decision Table with Domain Ontologies," *ICITM*, pp. 14–22, 2007.

12. C. Bizer, T. Heath, and T. Berners-Lee, "Linked Data - The Story so Far," *International journal on semantic web and information systems*, vol. 5, no. 3, pp. 1–22, 2009.

13. T. Halpin and T. Morgan, *Information Modeling and Relational Databases*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2nd ed., 2008.

14. C. Debruyne and R. Meersman, "GOSPL: A Method and Tool for Fact-Oriented Hybrid Ontology Engineering," in *Proceedings of the 16th international conference on Advances in databases and information systems* (T. Morzy, T. Härder, and R. Wrembel, eds.), ADBIS '12, (Berlin, Heidelberg, Germany), pp. 153–166, Springer-Verlag, 2012.

15. Y. Tang, "Semantic Decision Tables - A New, Promising and Practical Way of Organizing Your Business Semantics with Existing Decision Making Tools," *LAP LAMBERT Academic Publishing AG & Co. KG, Saarbrücken*, 2010.

16. H. Neuhaus and M. Compton, "The Semantic Sensor Network Ontology," in *AGILE Workshop on Challenges in Geospatial Data Harmonisation, Hannover, Germany*, pp. 1–33, 2009.

Society and Human Development
Psychology, Politics, Sociology
and Education

New States in the World Order – A Comparative Study of Kosovo and Crimea

Elvina Jusufaj

PhD candidate, Faculty of Philology, University of Tirana
`Ejusufaj@yahoo.com`

Abstract. Self-determination of peoples is a basic principle of the United Nations Charter. To maintain the balance of the international system and to preserve it from anarchy, United Nations Security Council decides case by case on the recognition of new states. UNSC decision is based not only in the satisfaction of Montevideo criteria but also to be firmly consistent in exercising its primary responsibility for maintaining international peace and security. The recognition of new states by the UNSC is the most important step on the legitimacy of that process. Its decision is based primarily in the consensus of states involved in the conflict. Kosovo and Crimea are assessed differently by UNSC. The recognition of Kosovo as an independent state was undermined by the veto power while Crimea's referendum was overwhelmingly not recognized. The paper will study if Kosovo sets a negative precedent in exercising the right of self-determination, and if Crimea is a similar case, through analyzing the key moments in the creation of these new entities; referendums and international recognition. Essential distinctions are highlighted. Kosovo is widely acknowledged and accepted a *sui generis* case. Its independence came as result of a long monitored comprehensive process, as the final option for stability and peace in the region. Crimea is a different case of a new entity created by the use of force. The forced imposition to recognize Crimea and accept its annexation by Russian Federation puts in danger the existing international order and brings back the geopolitics of 19th century.

Keywords: *Security Council, Kosovo, Referendum, Crimea, International Order*

1   Introduction

On 16th March, 2014, Crimea held a referendum to determine its future; its separation from Ukraine, after 60 years with the status of the Autonomous Republic, and its reunification with the Russian Federation. The Russian presidential decree of recognizing Crimea as an independent state was followed by the signature and ratification

of the interstate treaty for reuniting the Crimean Republic with the Russian Federation, adding a new constituent unit in the Federation.[1]

This is the second case when the territories of other independent countries have been forcefully seized, after the peaceful dissolution of the former Soviet Union, which is considered by President Putin as '*a major geopolitical disaster of the century*'.[2] After years of very difficult transition, Russian Federation got its momentum to rise as a regional power. In 2008, Russian troops occupied the territories of Abkhazia and South Ossetia from Georgia. This action reflected dissatisfaction of Russian Federation that former countries of USSR, within the zone of their influence, had expressed their tendencies and aspirations to become members of NATO and EU.[3] The Crimean events might indicate that Georgia's conflict didn't occur by chance but signaled the beginning of a geo-strategic scenario for restoring the power of the former Russian Empire.[4] As Brzezinski has noted, '*Without Ukraine, Russia ceases to be a Eurasian empire*'.[5]

The annexation of Crimea by the Russian Federation is considered in violation of the sovereignty and territorial integrity of Ukraine.[6] Russian diplomacy, assuming that the international community will not remain indifferent, attempted to legitimize this action by making use of precedents of the past, drawing parallels with Kosovo. Russia claimed that there cannot be two standards in the imposed rules for the self-determination of peoples and in the implementation of this principle of the UN Charter.

Can a Kosovo-Crimea parallel be drawn?! The two cases look very similar based on the principal of self-determination. The right of self-determination is one the main principles, but not the purpose of the United Nations Charter. In addition to Article 1 (2), Article 55 of the UN Charter proves the legal character of the right of self-determination by describing this right as a principle, as opposed to merely a political

---

[1] http://english.pravda.ru/russia/politics/18-03-2014/127129-russia_crimea_unite-0/

[2] Putin, V (2005), *Annual Address to the Federal Assembly of the Russian Federation*, Accessed June 20, 2014,
http://archive.kremlin.ru/eng/speeches/2005/04/25/2031_type70029type82912_87086.shtml

[3] 'In an April 17 interview, Putin declared, "We have reached a point beyond which we cannot retreat." …..The Kremlin's determination to stop the advance of Euro-Atlantic institutions was signaled in a smaller campaign against Georgia in 2008.'
Roberts, A. C. (2014) The Czar of Brinkmanship, Foreign *Affairs*, N.p., 5 May. Accessed June 20, 2014. <http://www.foreignaffairs.com/articles/141390/cynthia-a-roberts/the-czar-of-brinkmanship>.

[4] 'Russia's invasion of this Ukrainian region is at once a replay and an escalation of tactics that the Kremlin has used for the past two decades to maintain its influence across the domains of the former Soviet Union.'
Mankoff J (2014), Russia's Latest Land Grab; How Putin Won Crimea and Lost Ukraine, *Foreign Affairs,* Vol 93, No 3, Accessed July 4, 2014
http://www.foreignaffairs.com/articles/141210/jeffrey-mankoff/russias-latest-land-grab.

[5] Brzezinski Z (1997) *The Grand Chessboard: American Primacy and its Geostrategic Imperatives*, Basic Books, New York, p. 56

[6] *Territorial integrity of Ukraine*, Resolution of UN General Assembly, document A/68/L.39, 24 March 2014

programme.[7] If the principle of self-determination is necessary to the creation of the new state, it is not sufficient for its legitimacy. The Friendly Relations Declaration of UN General Assembly, although not binding *per se*, defines the criteria to exercise the right to self-determination.[8] To maintain the balance of the international system and to preserve it from anarchy, United Nations Security Council is the only august body which decides case by case on the recognition of new states. UNSC decision is based not only in the satisfaction of Montevideo criteria but also to be firmly consistent in exercising its primary responsibility for maintaining international peace and security. The recognition of new states by the UNSC is the most important step on the legitimacy of that process. Its decision is based primarily in the consensus of states involved in the conflict.

The secession of Crimea is assessed through analyzing its referendum for independence using the precedent of Kosovo independence. In the case of Kosovo we examine the process of Kosovo independence, the role of United Nations from the very beginning of the conflict to the present day, and the capacity of Kosovo to enter into relations with other states, as one of Montevideo criteria.

- The referendum for independence

The referendum is the main legal act that defines the will of people to self-determination. To be a legitimate one, the referendum needs to be free and fair under international standards; otherwise it cannot constitute a basis in international law for the sought territorial change.[9] Importantly, holding a free and fair referendum is only a necessary, but not a sufficient condition for a territorial realignment to be accepted as lawful by international law.[10]

The referendum needs to respect some procedural criteria in order to be considered legitimate and in accordance with international law. Despite its conduct in compliance with international standards, it is necessary to comply with specific guidelines, including the proper timing for such a process[11] and above all it requires the consensus of the respective governmental authority that holds the political sovereignty. In addition

---

[7] Simma B (2002) *The Charter of the United Nations; A commentary*, Oxford University Press, 2nd edt., p. 49

[8] United Nations General Assembly, *Declaration on Principles of International Law concerning Friendly Relations and Co-operation among States in accordance with the Charter of the United Nations,* document A/RES/25/2625, 24 October 1970

[9] Peters, A (2014) *Sense and Nonsense of Territorial Referendums in Ukraine, and Why the 16 March Referendum in Crimea Does Not Justify Crimea's Alteration of Territorial Status under International Law*, EJIL Analysis,
Accessed June 21, 2014 http://www.ejiltalk.org/sense-and-nonsense-of-territorial-referendums-in-ukraine-and-why-the-16-march-referendum-in-crimea-does-not-justify-crimeas-alteration-of-territorial-status-under-international-law/

[10] Ibid. Peters, A (2014)

[11] See Council of Europe, Parliamentary Assembly Recommendation 1704 (2005), *Referendums:Toward Good practices in Europe*, 29 April 2005, and
Venice Commission, *Code of Good Practice on Referendums, Guidelines on the holding of referendums* (CDL-AD(2007)008Rev), 20 January 2009

to a democratic decision-making process on territorial questions, its transparency and monitoring from the international community and a free, transparent and democratic voting process are some of the conditions required to be internationally recognized. Some of the most important and arguably hard international legal standards regarding the referendum process are, peacefulness; universal, equal, free and secret suffrage; framework conditions of freedom of media and neutrality of the authorities; and an international referendum observation.[12]

The referendum held in the Crimea was announced 10 days before Election Day, in terms of an invasion by Russian forces and under weapons pressure. The declaration of independence, adopted on 10 March 2014 from the parliament of the Autonomous Republic of Crimea and the Sevastopol City Council - with 78 votes in favor from the overall 100 members of parliament - included the decision to hold a referendum six days later. The declaration of independence stated:

"We, the members of the parliament of the Autonomous Republic of Crimea and the Sevastopol City Council, with regard to the charter of the United Nations and a whole range of other international documents and taking into consideration the confirmation of the status of Kosovo by the United Nations International Court of Justice on July, 22, 2010, which says that unilateral declaration of independence by a part of the country doesn't violate any international norms, make this decision,"[13]

The referendum was a non-transparent process, held in the absence of international observers. The ballot paper had no meaningful alternatives. Voters had two choices: to support reunification of Crimea with Russia as a subject of the Russian Federation, or to support the restoration of the 1992 Crimean constitution and the status of Crimea as a part of Ukraine.[14] The Crimean constitution of 6 May 1992 offers wider competences to the parliament of Crimea, including the sovereign right to establish relations with other countries, thus Crimean voters had no option but joining Russia.[15]

The referendum in Crimea is held in full violation of the Ukraine constitution which clearly defines the succession of its territories through a popular referendum and with the approval of the sovereign authority. The constitution clearly spells out that '*regulatory legal acts of the Verkhovna Rada of the Autonomous Republic of Crimea and decisions of the Council of Ministers of the Autonomous Republic of Crimea shall not contradict the Constitution and laws of Ukraine and shall be adopted in*

---

[12] Peters, A (2014) *Sense and Nonsense of Territorial Referendums in Ukraine, and Why the 16 March Referendum in Crimea Does Not Justify Crimea's Alteration of Territorial Status under International Law*, EJIL Analysis, 16 April 2014,
Accessed June 21, 2014 http://www.ejiltalk.org/sense-and-nonsense-of-territorial-referendums-in-ukraine-and-why-the-16-march-referendum-in-crimea-does-not-justify-crimeas-alteration-of-territorial-status-under-international-law/

[13] http://rt.com/news/crimea-parliament-independence-ukraine-086/

[14] http://www.bbc.com/news/world-europe-26514797

[15] Balmforth, R (2014) No room for 'Nyet' in Ukraine's Crimea vote to join Russia, *Online Reuters Edition,* Accessed at http://www.reuters.com/article/2014/03/11/us-ukraine-crisis-referendum-idUSBREA2A1GR20140311

accordance with and in pursuance of the Constitution of Ukraine, laws of Ukraine, acts of the President of Ukraine and the Cabinet of Ministers of Ukraine' (Article 135).[16] Despite the fact that Ukrainian Constitution clearly recognizes the authority of the Autonomous Republic of Crimea to organize and hold local referendums (Article 138/2), it (Crimea) is an integral constituent part of Ukraine and shall resolve issues relegated to its authority within the frame of its reference, determined by the Constitution of Ukraine (Article 134). The constitution of Autonomous Republic of Crimea clearly spells out that the conformity of statutory acts of the Supreme Rada of the Autonomous Republic of Crimea to the Constitution of Ukraine shall be solved by the Constitutional Court of Ukraine, pursuant to the Constitution of Ukraine. It also acknowledges the right of the President of Ukraine, in case of non-conformity of any statutory acts of the Supreme Rada of the Autonomous Republic of Crimea to the Constitution of Ukraine and Ukrainian laws, to suspend the effects of such statutory acts of the Supreme Rada of the Autonomous Republic of Crimea with simultaneous recourse to the Constitutional Court of Ukraine regarding the constitutionality of such acts. [17] The Venice Commission also concluded that "*The Ukrainian constitution prohibits any local referendum which would alter the territory of Ukraine and that the decision to call a local referendum in Crimea is not covered by the authority devolved to the authorities of the Autonomous Republic of Crimes*".[18]

The referendum of independence in Crimea is held in absence of free public debate, and with credible reports of intimidation. The result of the referendum, 97 % in favor of the secession from Ukraine and reunification with Russia, is considerably extreme with the results of opinion polls conducted in February 2014 by Ukrainian Research Institute, indicating that only 41 % of Ukrainian voters supported the region's incorporation into Russia.[19]

At the same time, Crimea's referendum and its unification with the Russian Federation is a clear breach of international law because it violates the international obligations of Russia, treaties and memorandums where Russia is a party member and is committed to respect the sovereignty and territorial integrity of Ukraine. Russia acted

---

[16] Constitution of Ukraine, Title X, The Autonomous Republic of Crimea, Accessed at http://www.president.gov.ua/en/content/chapter01.html

[17] Law of Ukraine, On Approval of the Constitution of the Autonomous Republic of Crimea, The Bulletin of the Supreme Rada of Ukraine (VVR), 1999, No. 5-6, Art. 43, As amended by the Law of Ukraine dated 19 March 2009, numbered 1167-VI,
Accessed at http://www.rada.crimea.ua/en/bases-of-activity/konstituciya-ARK

[18] European Commission For Democracy Through Law (Venice Commission), *Opinion on "Whether the decision taken by the Supreme Council of the Autonomous Republic of Crimea in Ukraine to organize a referendum on becoming a constituent territory of the Russian Federation or restoring Crimea's 1992 constitution is compatible with constitutional principles*", Opinion no. 762/2014, para.15, Venice, 21March 2014

[19]UK government's response to points made by President Putin in his address to the Russian Parliament on 17 March 2014, https://www.gov.uk/government/news/in-response-to-president-putins-address-to-the-russian-parliament
http://www.pewresearch.org/fact-tank/2014/03/06/regional-polls-show-few-ukrainians-russians-want-a-united-single-state/

in violence of Budapest Memorandums on Security Assurances (1994) - based on which Russia together with the United Kingdom of Great Britain and Northern Ireland and the United States of America reaffirmed their commitment to Ukraine, and confirmed to respect the independence and sovereignty and the existing borders of Ukraine[20] - and the Treaty on Friendship, Cooperation, and Partnership between Ukraine and the Russian Federation (May 31, 1997).

The referendum of Crimea is held based on the right of self-determination of the people of Crimea. But the right to self-determination is closely related to some specific human rights, such as freedom of expression, freedom of association and the right to free and genuine elections and it comes to further the general diffusion of democracy (Hilpold, 2009).[21] In the case of Crimea, despite Russian claims, the situation in the ground demonstrates no evidence of suppression or abuse by Ukrainian authorities in exercising such rights by Crimean citizens.

The referendum for the secession of Crimea from Ukraine was immediately followed by its incorporation into the Russian Federation. The right of secession accepted by the international law needs some specific condition to be materialized, such as the occurrence of Human Rights abuses, endangering the prospects of survival of retrospective groups (Hilpold, 2009).[22] In the case of Crimea we cannot identify such preconditions that necessitated the right of people for self-determination, therefore its secession from Ukraine.

The case of Kosovo is quite different. Despite additional elements of a wider autonomy granted by the Constitution of 1974 of Federal Republic of Yugoslavia, Kosovo was not accepted and recognized as a constituting element of the Federation with the right of succession. Based on this legal point of view, Kosovo was not considered by the Badinter Commission as a separate entity entitled to secede from Yugoslavia and not included in its opinions, despite the overwhelming results of the referendum in favor of the independence from Former Yugoslavia in 1991,[23] similarly to other republics. The opinions of Badinter Commission were of utmost importance because they defined the criteria, the right of secession as well as the basis for discussion and compromise which was later achieved with the Dayton Agreement (1995). This agreement succeeded to end the bloody conflict in Bosnia and Herzegovina and to create a neutral and functional equilibrium between the parties in conflict, but didn't generate stability in the Balkans. The escalation of Kosovo conflict, because of the suppression caused to the Albanian civilian population by the Yugoslav authorities, raised awareness of the public opinion and the international community. In 1998 OSCE mission in Kosovo gathered undisputed facts on human rights and humanitarian law violations on a staggering scale, often committed with extreme and appalling

---

[20]*Budapest Memorandums on Security Assurances*, *UN Document A/49/765,* December 5, 1994, Accessed at

http://www.cfr.org/arms-control-disarmament-and-nonproliferation/budapest-memorandums-security-assurances-1994/p32484

[21] Hilpold, P (2009) The Kosovo Case and International Law: Looking for Applicable

Theories, *Chinese Journal of International Law,* (2009) 8 (1): 47-61

[22] Ibid.

[23] Malcolm, N (1998) *Kosovo; A short History*, New York University Press, NY, p. 347

violence.[24] The delicate and sensitive situation in the region and the escalation of the conflict constituted the basis for including the case of Kosovo in the agenda of the United Nations Security Council, through adopting several resolutions and taking action based on Chapter VII of UN Charter considering "the situation in Kosovo as critical to peace and security in the region".[25] UN Security Council continued to monitor the situation in Kosovo, together with other international organizations (OSCE, NATO, etc) until the military intervention of NATO to prevent the spread of humanitarian catastrophe into the region. The resolution 1244 (1999) constitutes the ultimate UNSC resolution which gives an end to the humanitarian catastrophe in Kosovo caused by the holders of political sovereignty, and decides the creation of a United Nations Interim Administration Mission in Kosovo to monitor the transitional period toward the final political settlement of Kosovo.[26]

The presence of international institutions in Kosovo, since the end of the conflict in 1999, to administrate the province and free from Serbian political sovereignty, contributed towards the creation of a new multi-ethnic and democratic state. The option of unification with another country was never considered as possible by the international community. Furthermore, the 1991 referendum for independence has never been recognized as legitimate by the international community and accepted as the basis for the recognition of Kosovo's independence.

With reference to the advisory opinion of International Court of Justice on Kosovo (22 July 2010), which is used as a supporting argument in the declaration of independence of the Parliament of Autonomous Republic of Crimea and Sevastopol City Council, the Court found that the declaration of independence of Kosovo adopted on 17 February 2008 did not violate international law.[27] Kosovo, unlike Crimea - part of Ukraine and under its political sovereignty - is not under the sovereignty of Serbia, but under the international civil and security presence for more than a decade. UNSC resolution 1244 (1999) decided that the international civil presence, will facilitate a political process designed to determine Kosovo's future status, taking into account the Rambouillet accords (S/1999/648), and in a final stage, overseeing the transfer of authority from Kosovo's provisional institutions to institutions established under a political settlement.[28]

- The capacity to enter into relations with other states

The process of accepting Kosovo's independence went through several stages, which are assessed and supported by the international community. In terms of international

---

[24] OSCE Office for Democratic Institutions and Human Rights (1999) *KOSOVO / KOSOVA As Seen, As Told; The human rights findings of the OSCE Kosovo Verification Mission,*
Accessed at http://www.osce.org/odihr/17772?download=true

[25] UNSC resolutions 1160 (1998), 1199 (1998), 1203 (1998), 1239 (1998)

[26] United Nations Security Council 1244 (1999), 10 June 1999

[27] International Court of Justice, *Accordance with International Law of the Unilateral Declaration of Independence in Respect of Kosovo*, Advisory Opinion, I.C.J. Reports 2010, p. 403

[28] UNSC resolution 1244 (1999), 10 June 1999

relations, Kosovo is considered a *sui generis* case. The political solution of Kosovo came as a result of the humanitarian intervention by the international community to prevent the humanitarian catastrophe caused by repression and ethnic cleansing committed by Federal Yugoslav authorities against the Albanian civilian population of Kosovo.

Secondly, the case of Kosovo was monitored by the international community, the Security Council members and the Contact Group, which have defined the process and provided the necessary solution to ensure peace and stability in the region. This process is led and facilitated by the Contact Group, with three basic conditions to be respected: No return to Serbian sovereignty; No partition of Kosovo territories; No unification with other territories.[29]

These were the three basic negotiating principles which guided efforts of President Ahtisaari, Special Envoy of the Secretary-General of the United Nations, in resolving the political status of Kosovo. The Comprehensive Proposal for the Kosovo Status Settlement, which offered the best viable long-lasting solution for Kosovo and stability of the region,[30] was threatened by veto from the Russian Federation. The proposal got the support of all other members of the Security Council as a fair and proper solution. Today, Kosovo has been recognized by 107 states[31] and is in the process of Stabilization and Association with European Union[32]. The implementation of the First Agreement of principles governing the normalization of relations between Kosovo and Serbia, with the intermediate of Baroness Ashton, will lead towards stability in the region and creates future premises to promoting regional cooperation and peace in the Balkans.[33]

---

[29] "The settlement of Kosovo's status should strengthen regional security and stability. Thus, it will ensure that Kosovo does not return to the pre-March 1999 situation. Any solution that is unilateral or results from the use of force would be unacceptable. There will be no changes in the current territory of Kosovo, i.e. no partition of Kosovo and no union of Kosovo with any country or part of any country. The territorial integrity and internal stability of regional neighbors will be fully respected."
Guiding principles of the Contact Group for a settlement of the status of Kosovo, Security Council document, S/2005/709, 10 November 2005

[30] *Comprehensive Proposal for the Kosovo Status Settlement*, UN Security Council Document, S/2007/168/Add.1, 26 March 2007

[31] As of July 2, 2014, Togo is the 107th state to recognize Kosovo as an independent and sovereign state. Accessed July 3, 2014 http://www.mfa-ks.net/?page=2,4,2381

[32] Stabilization and Association Agreement negotiations successfully completed, 02/05/2014 Accessed July 1, 2014
http://eeas.europa.eu/delegations/kosovo/press_corner/all_news/news/2014/20140502_03_en.htm

[33] "It is a milestone that officials hope will enhance stability in the region and clear a path for both countries to join the European Union."
Dan, B (2013), Serbia and Kosovo Reach Agreement on Power-Sharing, *New York Times,* 19 April, Accessed July 1, 2014
http://www.nytimes.com/2013/04/20/world/europe/serbia-and-kosovo-reach-milestone-deal.html?_r=0
Gvosdev, K. N (2013) Kosovo and Serbia Make a Deal; Debalkanizing the Balkans, *Foreign Affairs,* April 24, Accessed July 1, 2014

The referendum for independence and Crimea's right to return to the Russian Federation was rejected by 13 members of the Security Council and abstained by China. In absence of a UN Security Council resolution to block the referendum of Crimea, the UN General Assembly adopted - with 100 votes in favor, 11 against and 58 abstentions - the resolution "The territorial integrity of Ukraine", recognizing the commitment of UN Member States to the sovereignty, political independence and territorial integrity of Ukraine, within its internationally recognized boundaries. The resolution emphasizes that '*the referendum held in the Autonomous Republic of Crimea and Sevastopol city on 16th March 2014, is not valid and cannot be the basis for changing the status of the Republic*'. It also demands from the countries, international organizations and specialized agencies not to recognize any alteration of the status of this republic and of Sevastopol city, based on the referendum "*and to refrain from any action or dealing that might be interpreted as recognizing any such altered status.*"[34]

The second argument used by the Russian Federation for reuniting with Crimea is the historical one. Crimea was part of Russia from 1783, when the Tsarist Empire annexed it a decade after defeating Ottoman forces in the Battle of Kozludzha, until 1954, when the Soviet government transferred Crimea from the Russian Soviet Federation of Socialist Republics (RSFSR) to the Ukrainian Soviet Socialist Republic (UkrSSR).[35] 56% of population in Crimea is Russian, 27 % Ukrainian and the rest is Tartar minority. Crimea is historically as close to Russia as to Ukraine. Although history serves to throw light on and understand current conditions, it is not the essential element to be considered in the developments of international relations and world system. The old dictum that international law is not an instrument for the revision of history, finds here a particular specification. Countries are defined in the course of historical and political developments based on respective agreements of the international order of the time. Without entering into specifications, many countries in Europe have territorial disputes with their neighbors. The affiliation of territories is not defined only by ethnicity, but rather by historical developments and international relations of the time, that cannot be undone. Otherwise the international order turns into a full anarchy. The historical facts are essential in defining and understanding recent developments, but they are not sufficiently enough. In the case of Kosovo, its unification with Albania has never been raised, even though it may be considered an indisputable historical right.

<http://www.foreignaffairs.com/articles/139346/nikolas-k-gvosdev/kosovo-and-serbia-make-a-deal>.

[34] *Territorial integrity of Ukraine*, Resolution of UN General Assembly A/68/L.39, 24 March 2014

[35] Kramer, M (2014) *The Transfer of Crimea from Soviet Russia to Soviet Ukraine, 1954*, http://wilsoncenter.org/publication/why-did-russia-give-away-crimea-sixty-years-ago#intro
"Meeting of the Presidium of the Supreme Soviet of the Union of Soviet Socialist Republics" March 19, 1954, History and Public Policy Program Digital Archive, GARF. F.7523. Op.57. D.963, L1-10. Published in Istoricheskii arkhiv," issue 1, vol. 1 (1992). Translated for CWIHP by Gary Goldberg. http://digitalarchive.wilsoncenter.org/document/119638

Finally, mentioning Kosovo as a precedent for the case of Crimea is an elusive case. It seems that is not used as a legal argument and a right precedent to justify its separation from Ukraine and annexation by Russia, but as a direct pressure on Europe and the United States in order to remind the international community that Russia is a world power and has the right to act on its zones of influence. Despite that, the Russian position on Kosovo is not consistent. In the Security Council meeting, the day after the declaration of Kosovo's independence, the Russian Ambassador stated that "The unilateral declaration of independence and its recognition are not in alignment with the provisions of the Helsinki Final Act, which clearly define the principles of territorial sovereignty and integrity of the states".[36]

The reference and the parallel set between Kosovo and Crimea demonstrates President Putin's concept of a world that must be again subject to the power of force rather than power of international law - a world that must operate based on consensus.

2   Conclusions

The paper highlights essential distinctions in assessing Kosovo and Crimea cases. The first distinction consists in the processes of the right of the peoples of Kosovo and Crimea for self-determination. Kosovo didn't declare the independence as a result of its expressed will for self-determination through the referendum. Independence referendums are additional elements in gaining independence and international recognition, where the basic element is the humanitarian intervention of the international community to stop the ethnic cleansing and to protect fundamental freedoms and human rights. The political solution of the Kosovo conflict came as the last viable alternative to secure peace and stability in the region. In the case of Crimea, the referendum is a strained act to impose a political solution based on force.

The second distinction consists in the fact that Kosovo crisis is widely discussed in the Security Council, which is never expressed against a final political settlement based on the will of people. The SC resolution 1244 (1999) includes, as its annex, The Rambouillet Accord that takes into account the political solution of Kosovo. Regarding Crimea, UNSC member states preliminary articulated the non-recognition of the referendum and strongly stated in favor of the territorial integrity of Ukraine.[37]

The third distinction arises from the political reasons behind the two cases. In the case of Kosovo, there was no influence from Albania in Kosovo's succession and the idea of national unification is always ruled out. In the case of Crimea, the influence of Russian Federation is of crucial importance and the annexation of its territories was part of the political solution.

---

[36] UK government's response to points made by President Putin in his address to the Russian Parliament on 17 March 2014, https://www.gov.uk/government/news/in-response-to-president-putins-address-to-the-russian-parliament

[37] UN Security Council action on Crimea referendum blocked, UN Security Council meeting, 15 March 2014, http://www.un.org/apps/news/story.asp?NewsID=47362#.U3D9soGSzVF

The efforts of the Russian Federation to find and apply the precedent of Kosovo in the case of Crimea are used as a precondition to legitimize the case in the UN Security Council. Russian Federation tried to highlight that the action of international community and the West is based on double standards. In fact, this argument exposes the negative influence of the veto power in the UNSC, when the veto of a sole country defines the position of the international body, entrusted to maintain international peace and security. At the same time, it demonstrates the rising tendencies of regional powers for zones of influence and a come-back of geopolitics and use of force, in an environment where liberal international order is based on consensus.

References

Simma, B (2002) *The Charter of the United Nations; A commentary*, 2ⁿᵈ Ed, vol I, Oxford University Press, New York

Malcolm, N (1998) *Kosovo: A short history*, New York University Press, New York

Brzeznisky, Z (1997) *The Grand Chessboard: American Primacy and its Geostrategic Imperatives,* Basic Books, New York

Roberts, A. C. (2014) The Czar of Brinkmanship, Foreign *Affairs*, Vol 93, No 3

Mankoff J (2014), Russia's Latest Land Grab; How Putin Won Crimea and Lost Ukraine, *Foreign Affairs,* Vol 93, No 3

Hilpold, P (2009) The Kosovo Case and International Law: Looking for Applicable Theories, *Chinese Journal of International Law,* (2009) 8 (1): 47-61

Peters, A (2014) *Sense and Nonsense of Territorial Referendums in Ukraine, and Why the 16 March Referendum in Crimea Does Not Justify Crimea's Alteration of Territorial Status under International Law*, EJIL Analysis

http://english.pravda.ru/russia/politics/18-03-2014/127129-russia_crimea_unite-0/

Putin, V (2005), *Annual Address to the Federal Assembly of the Russian Federation*, 2014, Kremlin, Moscow

http://rt.com/news/crimea-parliament-independence-ukraine-086/

http://www.bbc.com/news/world-europe-26514797

Richard Balmforth, R (2014) *No room for 'Nyet' in Ukraine's Crimea vote to join Russia*,   http://www.reuters.com/article/2014/03/11/us-ukraine-crisis-referendum-idUSBREA2A1GR20140311

Constitution of Ukraine, Title X, The Autonomous Republic of Crimea, http://www.president.gov.ua/en/content/chapter01.html

Law of Ukraine, On Approval of the Constitution of the Autonomous Republic of Crimea, The

Bulletin of the Supreme Rada of Ukraine (VVR), 1999, No. 5-6, Art. 43, As amended by the Law of Ukraine dated 19 March 2009, numbered 1167-VI

http://www.rada.crimea.ua/en/bases-of-activity/konstituciya-ARK

http://www.pewresearch.org/fact-tank/2014/03/06/regional-polls-show-few-ukrainians-russians-want-a-united-single-state/

Budapest Memorandums on Security Assurances, *UN Document A/49/765,* December 5, 1994,http://www.cfr.org/arms-control-disarmament-and-nonproliferation/budapest-memorandums-security-assurances-1994/p32484

UNSC resolutions, 1160 (1998), 1199 (1998), 1203 (1998), 1239 (1998)

United Nations Security Council 1244 (1999), 10 June 1999

International Court of Justice, *Accordance with International Law of the Unilateral Declaration of Independence in Respect of Kosovo*, Advisory Opinion, I.C.J. Reports 2010, 22 July 2010

Guiding principles of the Contact Group for a settlement of the status of Kosovo, Security Council document, S/2005/709, 10 November 2005

Comprehensive Proposal for the Kosovo Status Settlement, UN Security Council Document, S/2007/168/Add.1, 26 March 2007

Territorial integrity of Ukraine, Resolution of UN General Assembly A/68/L.39, 24 March 2014

Kramer, M (2014) *The Transfer of Crimea from Soviet Russia to Soviet Ukraine, 1954*, Wilson Center

UK government's response to points made by President Putin in his address to the Russian Parliament on 17 March 2014, 21 March 2014

UN Security Council action on Crimea referendum blocked, UN Security Council meeting, 15 March 2014

United Nations General Assembly, *Declaration on Principles of International Law concerning Friendly Relations and Co-operation among States in accordance with the Charter of the United Nations,* document A/RES/25/2625, 24 October 1970

Council of Europe, Parliamentary Assembly Recommendation 1704 (2005), *Referendums:Toward Good practices in Europe*, 29 April 2005

Venice Commission, *Code of Good Practice on Referendums, Guidelines on the holding of referendums* (CDL-AD(2007)008Rev), 20 January 2009

European Commission For Democracy Through Law (Venice Commission), *Opinion on "Whether the decision taken by the Supreme Council of the Autonomous Republic of Crimea in Ukraine to organize a referendum on becoming a constituent territory of the Russian Federation or restoring Crimea's 1992 constitution is compatible with constitutional principles*", Opinion no. 762/2014, para.15, Venice, 21March 2014

OSCE Office for Democratic Institutions and Human Rights (1999) *KOSOVO / KOSOVA as Seen, as Told; the human rights findings of the OSCE Kosovo Verification Mission*

Dan, B (2013), Serbia and Kosovo Reach Agreement on Power-Sharing, *New York Times*

Gvosdev, K. N (2013) Kosovo and Serbia Make a Deal; Debalkanizing the Balkans, *Foreign Affairs*

Voters' profiles and clustering. Factors that influence behavior

Evangelia N. Markaki[1] and Theodore Chadjipantelis[1]

[1]Department of Political Science, Aristotle University of Thessaloniki
markakievi@yahoo.gr

Abstract. With the present study we examine how and in which level some factors influence the creation of voters' profiles during the formation of political preferences People act in an environment with social, personal as well as professional networks where they interact, they develop activities, they undertake roles and they react. The influence that the network can exert on its members is also combined with external factors as networks are not only interactive but also interrelated or interconnected. In networks there are persons who influence more the others and others who are most influenced while forming their political preference. We examine the factors that influence political preferences in networks. The data collection took place in Greece. 1.103 persons participated. For the data analysis we used the ACP and Cluster Classification. From the results we see that 35% consider the personal interest as the crucial factor while forming political preferences. 33% evaluate more the position of the political parties on social, financial and national issues, 11% form political preferences assessing the positions of the political parties as well as the personal interest and finally 7.5% is influenced by the environment that means by personal, professional and social networks while forming political preferences. The present study focuses on the "new" science of networks connecting networks and factors that influence the formation of political preferences.

Keywords: *voting behavior, political preferences, ACP, cluster classification, networks*

1   Introduction

The political though as well as the formation of political preferences are complicated processes which consist of different stages but that are directly connected to interactions in the different spheres of action in which we live. (Christakis & Fowler, 2009). These interactions are part of the different spheres of actions where we live and which are typeset by networks. These spheres are the following: firstly, there is the sphere of interpersonal and familial relations that includes people with whom we live together or people who are our family (partners, parents and children). Secondly, there is the professional environment that includes the contact with colleagues. The third sphere of action is the social life that includes our friends and people who belong to our social environment.

In these spheres we shape our networks, we participate in networks and we interact with other people. The network is defined as a sum of people that are connected with each other and that can have one or more common characteristics (Christakis & Fowler 2009). Scientists believe that people create their networks but they are also influenced by others even if they do not know them (Christakis & Fowler 2009).

People belong in networks and influence each other (Christakis & Fowler, 2009). The influence in political behavior is related to different characteristics and factors that can be found in people in networks. We are units with many votes and we exert influence consciously or unconsciously (Christakis & Fowler, 2009). We are the mediators of different factors such as political knowledge, ideology, political identification etc. in our own networks, we shape them, we grow them but the networks affect us and mediate in the formation of political preferences (Christakis & Fowler, 2009; Inglehart & Norris, 2000; Weatherford, 1982).

2   Formation of political preferences: factors and influences

For the examination of the formation of political preferences studies use mostly multiple regression analysis with different variables such as the involvement in politics, the evaluation of the network as a place of interaction, ideology, the voter's social, personal and professional environment and the political identification. In the modern democracy where political parties participate with a dynamic way in the formation of political preferences via their political personnel, their political program, the regulations and their ideology, political identification seems to be the most important characteristic that is build via (Lenk, 1982; Miller, 1991; Huckfeldt & Sprague, 1987):

• The identification with the political leading figure of the party.

• The identification with the political program of the party (Calhoun Brown, 1981; Weatherford, 1982; Harris, 1994; Verba et al., 1995; Leighley, 1996; Radcliff & Davis, 2000; Putnam, 2000).

• The contact with political persons and political personnel (Berelson, Lazarsfeld, & McPhee, 1954). The interaction or the contact with political persons or the political personnel can be formal or informal (McClurg, 2003). The informal interaction is considered to be very important because exposes people to different messages mostly in a simple and unconscious way. The participation in networks makes people develop their political knowledge, be familiarized to politics and thus the formation of political preferences seems to happen in an easy way (Verba et al., 1995).

The participation in social networks such as in citizens' groups, in professional networks s but also the involvement in personal networks promotes the participation in politics, the political motivation as well as the political influence (Putnam, 2000; Leighley, 1996). Different characteristics influence the different voter's profiles (Weatherford,

1982). These characteristics can be the political knowledge, the political identification with the positions, the leading figure, the program of a political party, the contact with political persons or the political personnel and the political interest that needs monitoring of politics and of the public affairs.

The interaction in social and personal networks exposes people to political information in a different way than in professional networks (Mutz, 2002a, 2002b; Huckfeldt 2001; Giles & Dantico, 1982). Firstly, it is more flexible and secondly it does not conflict directly with other relationships, participation in networks or influences (McClurg, 2003).

The political behavior and the political preferences are connected to political parties, political personnel as well as to personal interest. The contact with political parties can help someone find a job, have opportunities for studies or make new friendships. These political contacts have also different projections: they organize and politicize networks with voters and supporters (Huckfeldt & Sprague, 1987; Zuckerman, 1994 ; Coffé & Need, 2010). In Greece political preferences are closely associated with the 'redemption' that is why our political system is characterized fundamentally clientelistic (Charalampis, 1989). It is already referred in previous studies that the younger voters choose taking into consideration the satisfaction of their personal interests and thus they form their political preferences at the end of the pre- elections campaign (Apospori, Ayloniti & Zisouli 2006). According to John Wilson and Robert E. Lane (Kouvertaris, 1997) the political preferences are connected not only to rewards but also to the voter's personal interest. Additionally Steven G. Rosentstone and John M. Hansen (Kouvertaris, 1997) support that people decide easier about their political preferences when these preferences give them specific benefits.

The retreat of the "left-right" antagonism, as is was illustrated by the continuous interchange of the two major parties (PASOK and Nea Dimokratia) in office (1974-1981 Nea Dimokratia, 1981-1989 PASOK, 1990-1993 Nea Dimokratia, 1993-2004 PASOK, 2004-2009 Nea Dimokratia, 2009-2011 PASOK) shows that today the importance of the axe "left-right" as well as the political identification has reduced. and this "left-right" antithesis gradually evolved as a "government-opposition" antithesis, without any ideological terms (Chadjipantelis, 2014). There is also the perception of the crisis from the political personnel. For the major part of the electorate, the economic crisis has its roots to the diachronic biased decisions and clientelistic structures of the State formation itself, which proves accurate what Anthopoulos (2007) states in his work. Consequently, the economic crisis is actually a crisis of the quality of Democracy, which does not pertain to the economy but to the characteristics of governance. In addition to that, the attention of citizens is moved away from welfare issues towards issues that pertain to the quality of Democracy. (Liu et al., 1998) The two major parties that interchanged in the government over the previous years received most of the citizens' disappointment. In many cases, decisions of the previous period were questioned.

Another important factor is the connection of the political party to the social action so as to deal with social problems (Huckfeldt & Sprague, 1987; Frederick, 1994). The socially oriented political parties gain more supporters (Huckfeldt & Sprague, 1987; Markaki, Sakas, & Chadjipantelis, 2011).

3   Methodology

For the analysis we use a two step procedure, computing firstly, via multivariate correspondence analysis, principal axes and loadings and secondly, through cluster analysis, the attitudes are grouped in clusters. Through this analysis, specific axes emerged, describing the data in less dimensions.

Correspondance analysis is a technique for displaying the rows and columns of a data matrix as points in dual low dimensional space. Furthermore analysing the data in a low dimensional space we are able to reveal data patterns finding groups of data points. This is a two-step procedure. In the first step we analyze the data using X square metric expressing the data using the derived axes. In the second step using hierarchical clustering we define groups of points assigning attitudes and so defining typologies of behaviour among people.

The analysis is implemented through the use of two-way cross tabulation, contingency tables, and correspondence analysis by using the pioneer program "M.A.D." [Méthodes de l'Analyse des Données], developed by Prof. Dimitrios Karapistolis. In "M.A.D." [Méthodes de l'Analyse des Données], Prof. Karapistolis (2010) has integrated and implemented digitally an abundance of methods created manually by the distinguished mathematician Jean-Paul Benzécri (1992).In order to identify specific attitudes of the population and form clusters that showcase the same attitude, we used the data of a study of 2011.

Consequently, data were analyzed using hierarchical cluster analysis in order to form homogenous groups of attitudes.

2.1 Data Collection

The research took place in Greece. The sample was 1.103 participants from urban and rural regions. The research took place from January to June 2011. The average response time to the questionnaire was approximately 20 minutes. The data collection was not random but with dispersion of questionnaires so as to represent each category of demographic characteristics examined in the research.

4   Results

The participants of the survey were asked to evaluate the importance of different factors for the formation of political preferences. The relevant frequencies are presented in the table below:

Table 1: The relevant frequencies (%) of the factors that influence political preferences

Factors	Little Important	Relevant Important	Medium Important	Quite Important	Very Important
Ideology	5.3%	9.0%	9.9%	33.7%	42.1%
Familial environment	20.3%	20.1%	22.9%	27.4%	9.3%
Social Environment	31.8%	29.6%	27.1%	9.6%	1.9%
Professional Environment	52.1%	24.4%	16.1%	5.7%	1.7%
Political Program	26.4%	18.9%	19.1%	25.4%	10.3%
Leading Figure of the political party	25.3%	15.0%	18.2%	25.1%	16.4%
Position of parties on national issues	14.2%	12.9%	16.8%	29.8%	26.3%
Position of parties on financial issues	11.0%	10.0%	14.4%	31.0%	33.7%

Position of parties on social issues (Health, Education)	9.6%	8.5%	12.4%	29.8%	39.7%
Personal Interest	31.0%	17.6%	16.9%	16.9%	17.7%
Contact with political persons or political personnel	49.7%	18.2%	14.8%	11.5%	5.7%
Familial tradition	44.2%	17.3%	15.3%	12.7%	10.5%

From the results of the table 1 we see that more than the 70% of the respondents believe that ideology can influence the formation of political preferences. In each case people vote today less ideologically oriented than in the past because in the past ideology was based on social classes and this has changed (Donald, Palmquist, & Schickler, 2002; Dodos, 2011). Social classes were connected to class identity According to Lipset (as mentioned in Kouvertaris, 1997) the political dimension of social classes changed because many different factors such as mobility and lack of social cohesion have contributed to this phenomenon. A typical example is that of the labor class which is no longer represented by one and only political party.

Even if the political reality, the ideological and political criteria have changed, the context of the political parties is different, the ideological differentiation is smaller and more difficult to be observed, ideology seems to motivate and prioritize the political behavior. (Kouvertaris, 1997). As far as it concerns the three basic spheres of action we observe that almost 40% of the respondents understand the influence that can have the familial environment on the formation of political preferences. On the other side 40% of the respondents believe that they are not influenced by their family on the formation of the political preference. The social as well as the professional environment are not considered to be influential for the formation of the political behavior. The importance of the family for the formation of political behavior has already be investigated because it has specific characteristics such as frequency, trust, intensity and intimacy characteristics that intensify the influence on political behavior. Additionally, influences on political preferences happen in a unconscious and not always evident way.

As far as it concerns the importance of the program as well as the leading figure of the political party 40% of the respondents believe that these are as not important factors for the formation of the political preference. We also see that 50% of the respondents believe that the way the political parties deal with social, national and financial issues is a very important factor for the formation of political preferences. This may happen because the way politicians deal with these issues influence the everyday life of citizens.

As far as it concerns the personal interest 30% of the respondents believe it plays an essential role for the formation of political behaviors. Many political preferences are based on personal interest that is why our political system is considered as clientelistic (Χαραλάμπης, 1989). People vote today less ideologically oriented. Thus, they form their political preferences trying to ameliorate their life and this is a clearly politically motivated decision.

On the other hand there are voters that they believe that their personal interest is not an important factor for the formation of the political preferences (50% of the respondents). During the last period the voters do not trust anymore the clientelistc political system (Sefertzis and Tsimogiannis, political communication experts. Personal contact, 2007).

As far as it concerns the contact between the voters and the politicians or the political personnel 70% of the respondents believe that this is not important factor for the formation of the political preference. This probably depicts the negative impression of a contact that could represent the propaganda or the political pressure towards voters (Charalampis, 1989).

About the familial tradition as factor for the formation of the political preference 70% of the respondents believe that they are not influenced by the political behavior of their family that as we have already said can happen in an unconscious way.

3.1 Results ACP and Hierarchical Clustering

Using the variables in the table below that represent some factors that influence the formation of political preferences we proceeded in ACP and Hierarchical Clustering.

Table 2: The variables- factors that influence political preferences

Q1	Ideology
Q2	Familial environment
Q3	Social Environment
Q4	Professional Environment
Q5	Political Program
Q6	Leading Figure of the political party
Q7	Position of parties on national issues
Q8	Position of parties on financial issues
Q9	Position of parties on social issues (Health, Education)
Q10	Personal Interest
Q11	Contact with political persons or political personnel
Q12	Familial tradition

In the table below we see how the factors create groups of factors after ACP and Cluster Classification.

Table 3: Croups and Classification

	Factors	Groups						Croups (Classification after ACP)
		I4	I5	I6	I3	I1	I2	
Q1	Ideology			SX			PP	**1**
Q2	Familial environment	EE		EE	EE			**3**
Q3	Social Environment			AR				**4**
Q4	Professional Environment			SX				**4**
Q5	Political Program						AR	**1**
Q6	Leading Figure of the political party							**1**
Q7	Position of parties on national issues					AR	AR	**2**
Q8	Position of parties on financial issues					PP	AR	**2**
Q9	Position of parties on					PP	AR	**2**

social issues (Health, Education)							
Q10 Personal Interest	PP	EE			PP		**3**
Q11 Contact with political persons or political personnel	SX	EE			SX		**3**
Q12 Familial tradition	SX	EE		EE			**3**
Clusters	**100**	**269**	**78**	**151**	**114**	**353**	1065
	9,4%	**25,3%**	**7,3%**	**14,2%**	**10,7 %**	**33,1 %**	
Factors (ACP)							
1	36%	G1	G2				
2	17,60 %	G3	G4				
3	9,80 %						
4	7,60 %						

After ACP the groups that are created are:
The first group named G1 includes the political program, the leading figure and the ideology. The next group named G2 includes the positions for the national, financial and social issues. The next factors that constitute the third Group G3 are the personal interest, the contact with politicians or political personnel, and the familial tradition. Finally, the group G4 includes the familial, social and the professional environment.

From the results we observe an internal opposition between the two first groups G1 and G2 as well as between the G3 and G4 by the second axe (first feature). We observe also opposition by the third axe (second feature) between Group 3 and Group 4. After the hierarchical clustering we took the subjects from where emerged the following groups:

[I1&I2] [(I3)[(I4,I5),I6]]

In the Group I1 10% of the total consider that the positions of the political parties for national, financial and social issues are important enough and very important factors for the formation of political preferences. The personal interest and the contact with the politicians or the political personnel are quite important or very important factors for the formation of political preferences.

In the Group I2 33.1% of the total consider that the ideology, the political program and the position of the political parties for the national, financial and social issues are quite important factors for the formation of political preferences.

In the Group I3 the 14.2% of the total believe that the familial environment is a factor of medium importance for the formation of political preferences.

In the Group I4 the 10.4% of the total consider that the familial environment, the familial tradition and the contact with the politicians or the political personnel are factors of relevant importance for the formation of the political preference and the personal interest is a very important factor for the formation of the political behavior. In the Group I5% 25.3% of the total consider that the familial tradition, the personal interest and the contact with the politicians or the political personnel are factors of medium importance for the formation of the political preference.

In the Group I6 the 7.3% of the total believe that ideology constitute a factor of relevant importance, the familial environment is a factor of medium importance, the social environment is considered as a factor of relevant importance and the professional environment is believed to be a factor important enough for the formation of political preferences.

Then, we see how much each variable and which variables contribute to the creation of groups. How many people belong finally in each of these groups? As we observe from the Group I4 and I5 the 35% of the total consider the personal interest as an important factor for the formation of the political preferences. From the Group I2 we see that 33% of the total assess the positions of the political parties for the national, social and financial issues for the formation of the political preferences From the Group I1 we see that 11% of the total form the political preferences based on the positions of the political parties as well as the personal interest. Finally, from the Group I6 we see that the 7.5%

of the total is influenced by the environment (social, familial, professional) for the formation of political preferences.

5   Conclusion – Discussion

Before summing up the main conclusions of the study it is necessary to present some important weaknesses. As the preference order or the importance of the factors may change over time for each individual depending on type of election (local, national, EU level), the economic situation or other issues it is necessary to say that the data were collected on 2011 and the study refers to that period where crucial changes in the greek political scene were about to happen with the evolution of new political parties and political persons. The study does not include these characteristics in the research and the analysis.

The present study is part of a research that focuses mostly in networks interrelating different factors referred with the influences in networks. The present study focuses more on factors than in networks and thus there is no completed context of how influences happen in networks.

Connecting factors and networks for the creation of political preferences there are also issues concerning the rationality of the voter, the circumstances or the protest voting as well as the online social life and networking. These are very interesting issues that can be answered in a future research.

The political scene in Greece has changed after 2010. The research took place on 2011 where changes as well as the emersion of many new political parties for example To Potami were not incorporated to the present study that is referring to a previous period.

With the present study we examined how and in which level some factors influence the creation of voters' profiles during the formation of political preferences. People act in an environment with social, personal as well as professional networks where they interact, they develop activities, they undertake roles and they react. The influence that the network can exert on its members is also combined with external factors as networks are not only interactive but also interrelated or interconnected. We examined the factors that influence political preferences in networks. From the results we see that 35% consider the personal interest as the crucial factor while forming political preferences. 33% evaluate more the position of the political parties on social, financial and national issues, 11% form political preferences assessing the positions of the political parties as well as the personal interest and finally 7.5% is influenced by the environment that means by personal, professional and social networks while forming political preferences.

Correspondence Analysis helped us to display the rows and columns of a data matrix as points in dual low dimensional space and Cluster Analysis using hierarchical clustering defines groups of points assigning attitudes and so defining typologies of behaviour among people.

The importance and the implementation of the present study to the filed of the electoral studies is that the study describes the attitudes that are connected to political parties and relates the axes of confrontation of the political parties with the axes of confrontation of attitudes. Thus we can delineate cleavages such as age, religiosity, urbanization, gender.

The existence of networks by itself can not explain why and how networks can influence people's behavior (Weatherford, 1982). The influence in political behaviour is related to different characteristics of the network. In networks, there are connections and interactions as well as different factors that influence voting behaviour such as political knowledge which is produced, consumed and recycled via political discussion, political identification, the assessment of political issues, the environment itself, ideology, and political interest. These elements exert influence on political behaviour because people "distribute" them to others as they develop the political interest and involvement in politics (Markaki & Chadjipantelis 2013; Huckfeldt & Sprague, 1987; McClurg, 2006; Richey, 2008).

6   References

Apospori E., Avlonitis G., Zisouli M. (2006) Political culture and perception of "political marketing" tools: a cross-generational comparison. *Refereed paper presented to the 3rd International Political Marketing Conference, Cyprus College*

Berelson, B. R., Lazarsfeld, P. F., & McPhee, W. N. (1954) *Voting: a study of opinion formation in a presidential campaign.* Chicago: The University of Chicago Press.

Calhoun-Brown, A. (1996) African american churches and political mobilization: the psychological impact of organizational resources. *Journal of Politics, 58*(4), 935-53.

Chadjipantelis, Th. (2014) Parties, Candidates, Issues: The effect of cricis in Greece and Portugal, paper under review for publication.
Christakis, N., & Fowler, J.M. (2009) *Connected. The surprising power of our social networks and how they shape our lives*. Little, Brown and Co.

Dodos, D.(2011) *Modernity and voting rights,* Epikentro Publications
Frederick, H. C. (1994) Something within: religion as a mobi- lizer of african-american political activism. *Journal of Politics, 56*, 42-68.

Giles, M., & Dantico, M. K. (1982) Political participation and the neighborhood social context revisited. *American Journal of Political Science, 26,* 144–150.

Coffé, H., & Need A. (2010) Similarity in husbands and wives party family preference in the Netherlands. *Electoral Studies, 29*(2), 259 – 268.

Donald, G., Palmquist, B. & Schickler, E. (2002) *Partisans hearts and minds: political parties and the social identities of voters*. New Haven, CT: Yale University Press.

Karapistolis, D. (2010) *Méthodes de l'Analyse des Données,* Altintzis Publications Thessaloniki (in greek)
Anthopoulos, C. (2007) The Greek political system as party-cracy, *Journal of Administrative Law, 3, 381-396 (in greek)*
Benzécri, J-P (1992) *Correspondence Analysis Handbook*. New York:Marcel Dekker

Harris, F. C. (1994) Something within: religion as a mobilizer of african american political activism. *Journal of Politics*, *56*, 42-68.

Huckfeldt,R., & Sprague, J.(1987) Networks in Context: The Social Flow of Political Information. *The American Political Science Review*, *81*(4), 1197-1216. Published by: American Political Science Association.

Kourvetaris, G.A. (1997) *Political sociology. Structure and process*. Allyn and Bacon.

Leighley, J. E. (1996) Group membership and the mobilization of political participation. *Journal of Politics, 58*(2), 447-63.

Lenk, K. (1982) *Political Sociology*, Publishing Paratiritis

Liu, J., Ikeda, K., & Wilson, M. (1998) Interpersonal environment effects on political preferences: the 'middle path' for conceptualizing social structure in New Zealand and Japan. *Political Behavior, 20,* 183-212.

Markaki, E. N. & Chadjipantelis, T (2013) Information Management in Networks. Measuring its influence with a new approach, *Procedia – Social and Behavioral Sciences* Elsevier, 73, 310 -318.

Markaki, E.N., Sakas, D.P., & Chadjipantelis, T. (2011) New political communication practices: no budget event management. The new challenge. *International Conference for Integrated Information*, 73-77. Kos, Greece.

McClurg, S. (2006) The electoral relevance of political talk: examining disagreement and expertise. Effects in social networks on political participation. *American Journal of Political Science*, 50(3), 737-754.

McClurg, S. (2003) Social networks and political participation: the role of social interaction in explaining political participation. *Political Research Quarterly*, *56*(4), 449-464.

Miller, W. E. (1991) Party identification, realignment, and party voting: back to the basics. American Political Science Review, 85(2), 557-568.

Putnam, R. D. (2000) *Bowling alone: the collapse and revival of American community*. New York: Simon & Schuster. New York.

Radcliff, B., & Davis, P. (2000) Labor organization and electoral Participation in industrial democracies. *American Journal of Political Science, 44*, 132-41.

Richey, S. (2008) The autoregressive influence of social network political Knowledge on voting behaviour, *British Journal of Political Science*, 38(3), 527-542.

Verba, S., Schlozman, K. & Brady, H. (1995) *Voice and equality: civic voluntarism in American politics.* Cambridge, MA: Harvard University Press.

Weatherford, S. (1982) Interpersonal networks and political behavior. *American Journal of Political Science*, *26*, 117-43.

Zuckerman, M. (1994) *Behavioral expressions and biosocial bases of sensation seeking*. Cambridge University Press.

Charalampis, D. (1989) C*lientelism and populism. The extra-institutional consensus in the Greek political system.* Publishing Exantas.

Developing a Conceptual Framework for Detecting the Emergence of New Scientific Paradigms

Gorgiev Anka[1], Dr. Dimitriadis Nikolaos[2], Dr. Nikolaidis Dimitrios V.[2], and Dr. Martin Chris[3]

[1]South East European Research Centre (SEERC), Research Center of the International Faculty of The University of Sheffield, CITY College
`agorgiev@seerc.org`

[2]International Faculty of The University of Sheffield, CITY College
`ndimitriadis@city.academic.gr, d.nikolaidis@city.academic.gr`

[3]The University of Sheffield
`C.Martin@sheffield.ac.uk`

Abstract. In today's science, the word 'paradigm' has become overly used in numerous scientific disciplines. Even though its understanding was initialized by the scientific historian Thomas Kuhn more than fifty years ago, there is still no scientific method that suggests how the occurrence of a new paradigm can be confirmed. Having this in mind, the following literature-based paper attempts to build a framework which facilitates the identification of paradigm occurrence in a scientific manner. This research study suggests implementation of the grounded theory method in constructing a framework that will define the prerequisites for the acknowledgement of new paradigms. Through extensive analysis of historical data, concepts that are considered important for paradigm research will bewere extracted and grouped into meaningful categories, which together will form a comprehensive framework. Such approach will provide initial understanding of the detection of new paradigms as the phenomenon crucial for the progress of science.

Keywords: *Paradigm, Paradigm Shift, Theory-Building*

Introduction

Today's scientific landscape represents a result of numerous milestones that took place throughout the scientific discourse. The analyses of any distinctive chronological periods of any particular discipline, can provide an understanding of the particularities of the given period and how it contributed to its' contemporary content. Distinct chronological periods are often defined on the basis of a specific paradigm that was dominant during that period of time. For this reason, it is feasible to

investigate the importance of specific periods, or paradigms, to the development of knowledge for a particular discipline.

To date, the word paradigm has become overly used in numerous scientific disciplines (Feyeraberd, 2010). It can be found in a wide range of scientific fields, such as biology, medicine, psychology, sociology, politics, technology, law, management, as well as in areas such as religion, nutrition, environment, and even sports (Packowski and Francas, 2013; Quaid, 2012; Ivancevic, et al 2011; Ferguson, 2009; Malla, 2008; Schwarz, 1999; Tapscott and Caston, 1992; Morin, 1992; Ritzer, 1975). Yet, paradigms are usually acknowledged and recognized only after the new paradigms emerge, as there is currently no scientific way to identify the paradigm at the time of their occurrence. Having this in mind, a framework for paradigm identification might have very broad applicability and could significantly contribute to the understanding of how scientific knowledge develops. This study aims to build a framework which facilitates the identification of paradigm occurrences in a scientific manner.

Literature Review

The understanding of the concept behind the word 'paradigm' was initialized by the scientific historian Thomas Kuhn in his highly acknowledged book "The Structure of Scientific Revolutions" (Dear, 2012). According to Kuhn (1962), paradigm represents a set of habits that is well accepted by the scientific community at the given time. As such, paradigms outline the pattern by which scientific problems are being solved. In other words, he argues that a paradigm is defined by the scientific community in terms of the commitment its members make to the common problems and methods used to solve these problems.

Masterman (1965) claimed to have identified twenty one definitions of the notion of paradigm as Kuhn used it in his book. According to her interpretation, Kuhn has presented three different groups of contextual meanings of paradigm - metaphysical, sociological and contractual. The metaphysical meaning poses that paradigm represents a new way of viewing the world, as well as a framework that governs the perception of the world and the habits that are exhibited at the time of paradigm's ruling. This was advocated by Barker (1992), when he claims that paradigms act as 'physiological filters'. In this sense, scientists and practitioners view the world and everything that it represents through the prism of a paradigm, hence everything they perceive is influenced by the ruling paradigm. Lastly, the sociological meaning involves both the institutional involvement in paradigm governance and the scientific achievements that are being accumulated by the scientific community.

As argued by Ritzer (1975), politics play an important role in understanding, perceiving and embracing paradigm and its rules, since the power distribution towards one paradigm enables this paradigm to establish dominance over other paradigms. The contractual meaning includes the instrumentations, tools and textbooks supplied by the paradigm proponents in an effort to provide the platform for tacit and explicit knowledge creation. New scientists are introduced to the paradigm through the education system which is mainly based on the scientific achievements of the dominant one, the familiarization with the instruments and tools takes place while practicing the science and contributing to the empirical evidence (Nonaka, 1994). This was recently confirmed by Douglas (2012), who analyzed economic methodology used by Nobel laureates over time, realizing that there is a noticeable preference towards one paradigm, which can be perceived as a paradigm shift.

Although Thomas Kuhn is considered to be the founder of the notion of paradigm, there are other philosophers of science that have focused their work on conceptualizing the progress of science. What Kuhn called paradigm, Lakatos called research programme, while Laudan referred to it as research tradition (Laudan, 1978; Lakatos, 1965; Kuhn, 1962). According to Lakatos (1965), research programmes involve methodological rules that guide scientists towards directions that should be pursued or lead them away from those that should be avoided. He argued that research programmes represent a framework consisting of several individual theories, which as the whole constitute a unique belief system shared by a scientific community. Similarly, Laudan (1978) placed research traditions as an essential part in his view of the scientific process. According to him, research traditions are constituted of several individual theories and are defined by metaphysical and methodological guidelines that are specific only to a sole research tradition (Laudan, 1978). In addition, he argued that research traditions provide a set of guidelines for developing theories that solve the problems.

Even though the understanding of paradigms is heavily dependent on definitions and philosophical notions, scientists from diverse disciplines have embraced it in their works. Smith (1982) defines it as a set of assumptions that help scientists in understanding the world and predict its performance. Moreover, Harman (1979) adds that paradigms represent tacit understanding of the world and, as such, cannot be taught, but can only be transferred to future generations through experience. Similarly, Ferguson (1978) believes that paradigms provide a scheme for developing thoughts about the world and the reality surrounding it. Using a sports metaphor, Barker (1992) believes that paradigms provide rules that are needed in order to "play the game"; in other words, it defines the boundaries and gives instructions for how one should behave within those boundaries.

Paradigm Shift

Taking all the aforementioned definitions in consideration, one can claim that paradigm represents distinctive section in a discipline's history, and is characterized by its own set of rules and beliefs, which project a particular perception of the world. However, they seem to be appropriate only for a given time or context, which suggests certain changes after the rules of the paradigm are no longer applicable. In a study conducted by Hong, Hess and Hardin (2013) the authors managed to test which rules apply to which paradigm and utilized this information to increase the appropriateness of the paradigm applied to a given context. Kuhn (1962) referred to this as paradigm shift.

According to Kuhn's perception of the scientific revolution, there are three phases that occur in paradigm shifts (Kuhn, 1962). Firstly, a pre-paradigm period takes place, which is characterized by continuous debates over the validity of scientific problems, methods used and standards for their solution. This phase represents the period of crisis during which new schools of thought emerge and competing alternative paradigms are brought to light. However, sometimes during these activities, more anomalies are identified and isolated as the existing theory is not able to provide solutions for pending problems, yielding new theories brought forward to test. By finding promising contenders for the new ruling paradigm, science enters the second phase of the revolutionary science. At this stage, new problems are being identified and new methodologies are being tested; the interest of the scientific community is rising to the point of achieving critical mass; the previous anomaly has gathered sufficient empirical evidence as to be perceived as the new law. All this results into a new paradigm being established (Kuhn, 1962). As the final stage of the scientific revolution, a 'normal science' period takes place, where debates tend to be resolved and the new theory/paradigm is accepted (Kuhn, 1962).

Regardless which scientific discipline is in question, paradigm shifts seem to start in the same way, with the gap that has occurred between the actual observations and the predisposed expectations for such observation (Feyerabend, 1993). The term paradigm shift is generally used to depict a significant change (House, 2012). More so, the term Kuhnian paradigm shift seems to be used to indicate even greater change, one expected to bend the boundaries in the process of transformation (Achrol and Kotler, 2012). Barker (1992) argues that new paradigm catalysts can appear at different stages of the paradigm curve, and consequently the process of paradigm shift can have different trajectory and be subject to different levels of resistance to change. Henriksen (2013) adds that even though paradigm shifts are higher order changes, they can be facilitated by new models. So far, the process that takes place during the

paradigm shift , as identified by historians, philosophers, sociologists and practitioners of science, is the following: with the emergence of unsolved problems, scientists need to drop the existing practices and take a step back to consider the problems in more detail; after encountering the possible solution, paradigm shifters and pioneers engage in dissemination of the new idea with the goal to reach the critical mass; from that point on, the idea can diffuse in self-sustaining way (Rogers, 2003; Gladwell, 2002; Barker, 1992; Feyerabend, 1993).

Nevertheless, only one part of scientific progress represents the creation of the cognitive structures, such as concepts, theories, and paradigms; the other part of the science includes the social structures that facilitate the creation of knowledge through financial resources, such as institutions, societies and academies (Fuchs, 1993; Mendelsohn, Weingart and Whitley, 1977). This social aspect seems to be very important in the paradigm process, especially at stages that supersede the identification of the new idea and proceed to the stages of dissemination of the new idea (Faria, Besancenot and Novak, 2011).

Even though the interest and research surrounding paradigms is present for more than fifty years, it can be argued that there are still questions that have not been answered yet. This research study aims to contribute to the knowledge of paradigms and provide deeper understanding of the paradigm phenomenon. In particular, the study will address the question of what prerequisites need to be satisfied for a new emerging paradigm to be accepted.

Proposed Methodology

According to Pirson and Lawrence (2009), the emergence of widely accepted value models in any scientific discipline indicates a paradigm shift. However, for the new theory to be accepted as a paradigm shift, it requires more than just the acknowledgement of its occurrence. Interestingly enough, there is no clear methodology that can predict the occurrence of a new paradigm. On the contrary, there is only data that can point towards certain patterns in the history of science, thus from the analysis conducted by philosophers and historians of science. Yet, many new ideas have been rejected because they do not fit the rules of the ruling paradigm. Such circumstances are likely to result in a slower pace of scientific progress. By improving awareness of the new emerging belief systems, the academic community can facilitate the progress of scientific knowledge sustainably. Therefore, there is a need for the issue of paradigm identification to be explored to a fuller extent, particularly in an applied context (Brand, 2008).

There can be a continuous growth of knowledge resulting from consistent upgrades, which would allow for rapid and more consistent adoption of new findings.

At the same time, it could prevent dramatic shifts that cause strong resistance to change, as evident from the history of science. This can have effects in academia, through more informed understanding of the current context of scientific activity, as well as in the industry, through sustainable business strategies.

As Lynham (2002) suggests, the most appropriate way to generate a new framework is by implementing theory-building techniques. The advantage of this approach is that theory-building methods can enable the scientific community to "understand, explain, anticipate, know, and act in the world in better and more informed ways, and to better ends and outcomes" (Lynham, 2002, p.222). It has been argued in the past that theory-building as a research tool helps the process of knowledge growth in its basic sense (Markóczy and Deeds, 2009). In reality, the nature of theory-building research involves the interrelation of theory and practice, as well as the necessity for theory to be applicable in the real-life situations (Storberg-Walker, 2006).

In the case of the paradigm framework, it appears that the most appropriate strategy for theory building is deduction. As Lynham (2002) argues, the initial starting point should be the research and data collection that would lead to the formulation of the theory. In addition, this approach seems to be in line with the theory-building process described by Lynham (2002) and Goulding (2000); the authors suggest that the process of building the grounded theory involves the identification of the research area and data collection, interpretation of this data, theoretical sampling and finally concept development.

Following the suggestion of Goulding (2000), this study will employ the method of grounded theory development, throughout which it will address the origin of the early concepts and categories that have been identified in the area of paradigm research. Two major reasons behind such decision include the idea behind grounded theory, which essentially represents theory formulated from data, and the fact that scientific knowledge represents well-documented data set (Glaser and Strauss, 1999). The literature will be reviewed systematically in an effort to identify key themes mentioned in paradigm research, which will help in definition of concepts which will be further defined and then grouped in significant categories that together represent a mutually dependent construct. It will also elucidate the individual concepts within those categories in an effort to provide understanding of the theoretical construct as an independent entity. The result of this study will be provided through the framework for paradigm identification that will be constructed.

Conclusions

There is no doubt that paradigms represent very important concepts, both in the historical process of scientific knowledge and in contemporary scientific dynamics. The notion of paradigm itself also represents an important concept in the scientific landscape. However, there is still no clear scientific method that uncovers the process of paradigm identification, neither after nor before a specific paradigm shift takes place. In order get one step closer to this phenomenon, this research study will employ the grounded theory method in constructing a comprehensive framework that will provide deeper understanding of the paradigm dynamics as an important element in the progress of science.

References

Achrol, R. and Kotler, P. (2012) Frontiers of the Marketing Paradigm in the Third Millennium. *Journal of the Academy Of Marketing Science.* 40(1), 35-52.

Barker, J.A. (1992) *Paradigms: The Business of Discovering the Future*, HarperCollins Publishers, New York.

Brand, V. (2008) Empirical Business Ethics Research and Paradigm Analysis. *Journal of Business Ethics.* 86(4), 429-449.

Dear, P. (2012). Fifty years of Structure. *Social Studies of Science*. 42(3), 424-428.

Douglass, B. (2012) Economic Methodology and Nobel Laureates: Confirmation of a Methodological Paradigm Shift. *American Journal of Economics & Sociology*. 7(5), 1205-1218.

Faria, J., Besancenot, D. and Novak, A. J. (2011) Paradigm Depletion, Knowledge Production and Research Effort: Considering Thomas Kuhn's Ideas. *Metroeconomica*, 62(4), 587-604.

Ferguson, M. (2009) *The Aquarian Conspiracy: Personal and Social Transformation in Our Time*, Tarcher, Los Angeles.

Feyerabend, P. (1993) *Against Method*, Verso Books, New York.

Feyeraberd, P. (2010) *Against Method,* Verso, New York.

Fuchs, S. (1993) A Sociological Theory of Scientific Change. *Social Forces*. 71(4), 933-953.

Gladwell, M. (2002) *The Tipping Point: How Little Things Can Make a Big Difference*, Little, Brown and Company, New York.

Glaser, B.G. and Strauss, A.L. (1999) *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Aldine Transaction, New Jersey.

Goulding, C. (2000) Grounded Theory Methodology and Consumer Behaviour, Procedures, Practice and Pitfalls. *Advances in Consumer Research*. 27(1), 261-266.

Harman, W. (1979) *An Incomplete Guide to the Future*, W. W. Norton & Company, New York.

Henriksen, L.F. (2013) Economic Models as Devices of Policy Change: Policy Paradigms, Paradigm Shift, and Performativity. *Regulation & Governance.* 7 (4), 481-495.

Hong, W., Hess, T.J. and Hardin, A. (2013) When Filling the Wait Makes It Feel Longer: A Paradigm Shift Perspective For Managing Online Delay. *MIS Quarterly*. 37 (2), 383-406.

House, R. (2012). Fifty Years on From Kuhn: Towards an Urgently Needed Paradigmatic Transformation in Education. *Interconnections.* (8), 24-32.

Ivancevic, T.T., et al (2011) Paradigm Shift for Future Tennis: The Art of Tennis Physiology, Biomechanics and Psychology. *Cognitive Systems Monographs*. 12, p. 227-258.

Kuhn, T.S. (1962) *The Structure of Scientific Revolutions,* University of Chicago Press, Chicago.

Lakatos, I. (1965) Falsification and Methodology of Scientific Research Programmes, in Lakatos, I. & Musgrave, A. (eds), *Criticism and the Growth of Knowledge: Proceedings of the International Colloquium in the Philosophy of Science,* Cambridge University Press, London, 91-197.

Laudan, L. (1978) *Progress and Its Problems: Towards a Theory of Scientific Growth*, University of California Press, Ltd., Los Angeles.

Lynham, S.A. (2002) The General Method of Theory-Building Research in Applied Disciplines. *Advances in Developing Human Resources*. 4, 221-241.

Malla, K. (2008) Current State of the Law of International Watercourses: Progress and Paradigm Shifts 1815–2008. *Nordic Journal of International Law*. 77(4), 461-

Markóczy, L. and Deeds, D. L. (2009) Theory Building at the Intersection: Recipe for Impact or Road to Nowhere?. *Journal of Management Studies*. 46(6), 1076-1088.

Masterman, M. (1965) The Nature of a Paradigm, in Lakatos, I. & Musgrave, A. (eds), *Criticism and the Growth of Knowledge: Proceedings of the International Colloquium in the Philosophy of Science,* Cambridge University Press, London, 59-91.

Mendelsohn, E., Weingart, P. and Whitley, R. (1977) *The Social Production of Scientific Knowledge*, D. Reidel Publishing Company, Boston.

Morin, E. (1992) From the Concept of System to the Paradigm of Complexity. *Journal of Social and Evolutionary Systems.* 15(4), 371–385.

Nonaka, I. (1994) A Dynamic Theory of Organizational Knowledge Creation. *Organization Science.* 5(1), 14-37.

Packowski, J. and Francas, D. (2013). Lean SCM: A Paradigm Shift in Supply Chain Management. *Journal of Business Chemistry*. 10(3), 131-137.

Pirson, M.A. and Lawrence, P.R. (2009) Humanism in Business – Towards a Paradigm Shift?. *Journal of Business Ethics.* 93(4), 553-565.

Quaid, J. A. (2012) Making Sense of The Shift in Paradigm on Cartel Enforcement: The Case for Applying a Desert Perspective. *McGill Law Journal*. 58(1), 149-198.

Ritzer, G. (1975) Sociology: A Multiple Paradigm Science. *The American Sociologist.* 10, 156-167.

Rogers, E.M. (2003) *Diffusion of Innovations,* Free Press, New York.

Schwarz, C.A. (1999) *Paradigm Shift in the Church: How Natural Church Development Can Transform Theological Thinking*, Churchsmart Resources, Illinois.

Smith, A. (1982) *Power of Mind*, Simon & Schuster, New York.

Storberg-Walker, J. (2006) From Imagination to Application: Making the Case for the General Method of Theory-Building Research in Applied Disciplines. *Human Resource Development International*. 9(2), 227-259.

Tapscott, D. and Caston, A. (1993) *Paradigm Shift*, McGraw-Hill Company, New York.

Self Disgust validation in Greek population

Marianna Tsatali[1], Lambros Lazuras[2], Paul Overton[3] and Ana B. Vivas[2]

[1]South East European Research Centre, Psychological Studies
mtsatali@seerc.org

[2] International Faculty of the University of Sheffield, Psychology Department, City College
lazuras@sheffield.ac.uk, vivas@city.academic.gr

[3]The University of Sheffield
p.g.overton@sheffield.ac.uk

Abstract. The conceptual framework of self-disgust is described as the experience of disgust towards the self, which is considered to belong to the general category of self-conscious emotions. Self-disgust is under researched in relation to the other emotions, although it seems to play a crucial role in psychopathology. The present study aimed at validating the Self-Disgust Scale (SDS) in Greek population. The Greek version of the SDS was administrated with the following psychometric tools: the Beck Depression scale, the Rosenberg self-esteem scale, the Life orientation test, the Self description questionnaire and the Disgust scale. A total of 234 healthy adults (52 males and 182 females, aged from 18 to 32 years) completed the study. Results showed that the SDS has a very good internal consistency (Cronbach's α= .86) in the Greek population. Moreover, SDS had an overall satisfactory construct validity. That is, the translated version of the SDS appears to maintain the psychometric properties of the original English version, and thus is adequate to study Self-disgust in Greece.

Keywords: *Self-Disgust, Greek population, validation*

1   Introduction

Self-disgust can be identified as a strong negative feeling directed toward the self, which is part of one's personal belief system (Beck 1967). The concept of self-disgust is based not on whether the individual is conscious about the disgust elicitors (Rusch et al., 2011), the most significant prerequisite is that the elicitors are integrated into the self-concept, that they are something which is close to someone's personal sense of the self (Power & Dalgleish, 2008). Furthermore, self-disgust is related to aspects of the self which are significant and unchangeable for the individual's self image (Powell et al., in press).

The importance of attitude towards the self was suggested by Beck in his theory of the cognitive triad (Beck 1995). It has been found that depressive symptoms are strongly correlated with attitudes and subsequent emotional beliefs (Whelton & Greenberg, 2005). Specifically, distorted notions of the self, such as self-disgust, can be fundamental prerequisites for symptoms' appearance. Thus, studies have shown that this cognitive-affective construct of self-disgust is usually present among individuals with depressive symptomatology or clinical depression (Schienle et al., 2003). Additionally, according to literature (Overton et al., 2008), self-disgust is a core mediator between one's personal dysfunctional cognition and existing depressive symptoms, which highlights the significance of how one evaluates the self for the clinical domain.

Although recent studies support a significant role of self-disgust's in the etiology of depressive psychopathology (Power & Dalgleish) and other clinical disorders (Phillips et al., 1998), little is known about this emotion. Generally, self-disgust belongs to the category of self conscious emotions, and is regarded as a similar to, but distinct from, shame, dislike and guilt, although its separation from these emotions is often under recognized. For that reason it has to be identified independently from the above emotions and treated separately during therapeutic interventions and strategies.

For all the reasons mentioned before, dysfunctional responses towards the self have to be described and studied further to develop our understanding of the conceptual framework into which self-disgust fits and subsequently bridge the gap between self-disgust and the various psychopathologies in which self-disgust appears to play a role. To do this, current scientific research has to develop scales for self-disgust assessment for the purpose of understanding and evaluating this psychological characteristic (Olatunji et al., 2012).

In 2008, Overton and colleagues (Overton et al., 2008) created the only available self-report scale of self-disgust, the Self-disgust scale (SDS). This scale consists of 18 items and the sum score is calculated from 12 items, because there are 6 fillers, with a maximum score is 84 and the minimum is 12. Factor analysis has shown that the scale measures two latent constructs-disgust towards the 'self' and disgust towards one's behavior.

In order to advance our understanding of self-disgust, and potentially use it as a point of therapeutic intervention, not only do we need to develop scales to measure the construct, but we also need to translate those scales so that studies can be conducted worldwide, and incorporate important cross cultural investigations. To this end, we developed a Greek translation of the SDS.

2   Method

Participants

In the study we recruited 234 healthy adults (52 males and 182 females from 18 to 32 years old). The participants' demographic information (age, gender, marital status and educational level) is presented in Table 1. The sample included 234 undergraduate and postgraduate students from several public and private Universities in North Greece. Participants were recruited from researchers' personal contacts, professional networking and announcement boards.

Instruments

The Self-Disgust Scale (SDS) is 18-item psychometric tool (Overton et al., 2008). The sum score is calculated from 12 items, because there are 6 fillers, with a maximum score is 84 and the minimum is 12. Factor analysis has shown that the scale measures two latent constructs – disgust towards the 'self' and disgust towards one's behavior.

The translation-back translation method (Hambleton, 2001) was used to translate the English version of the SDS into Greek. Specifically, a group of Greek experts translated the original English version intro Greek. Then a group of bilingual researchers in the UK performed the back translation from English to Greek. The originally translated and the back-translated Greek versions were then compared for consistency, relevance and meaning of the content.

Before the scales' administration, participants filled out the demographics' questionnaire. After that, they were given the following psychometric scales: Disgust Sensitivity Scale (Rozin et al., 2008), Beck Depression Inventory (Fountoulakis et al., 2003), Life Orientation Test (Scheier et al., 1994), Self Description Questionnaire (Marsh, 1984) and Rosenberg's Self- Esteem Scale (Robins et al., 2001).

- **Disgust Sensitivity Scale (DSS)**

The Disgust Sensitivity Scale (DSS) (Haidt et al., 1994) is a psychometric tool which measures disgust sensitivity. According to the most recent analysis (Olatunji et al., 2008), it consists of three factors: Core disgust, animal reminder disgust, and contamination based disgust. DSS showed satisfactory statistical properties (Cronbach's a = .87 for the whole scale, .78 for core, .78 for animal reminder, and .54 for contamination) (van Overveld et al., 2011). It's the most popular scale for measuring disgust, and has been translated in many languages. The Greek translation of DSS was conducted by Sotiropoulou et al. (2012).

- **Beck Depression Inventory (BDI-II)**

BDI is a well known tool for depressive symptoms' investigation. The scale consists of 21 items (scored from 0-3). The scale does not include reverse items. The Greek version of the BDI (Fountoulakis et al., 2006) shows very good psychometric proper-

ties (Cronbach's alpha= .93, and test-retest with Pearson coefficient between .75 and .98). Hence, the scale has been adapted for the Greek population efficiently.

- **Rosenberg Self-Esteem Scale (RSES)**

Rosenberg Self-Esteem Scale (Rosenberg, 1965) is the most popular scale for measuring self-esteem. It includes 10 statements from strongly agree to strongly disagree. This psychometric tool has been translated into 28 languages (Schmitt & Allik, 2005) and its factor analysis is differentiated across nations. Rosenberg's scale in a Greek population has a Cronbach's alpha coefficient exceeding 0.70 (Koumi & Tsiantis, 2001) and hence has good psychometric properties (Koumi, 1994).

- **Life Orientation Test (LOT-R)**

The life orientation test (Scheier et al., 1994) is an easily administered psychometric tool which measures optimism. This two factor structure test (LOT-R-revised) has been validated in Greek population (Lyrakos et al., 2010) and shows satisfactory statistical properties (Cronbach's a = .71 and unitary structure and stability over a 3-months period (r = .66). In addition, its convergent validity is very good too (r = .73). As a result, it can be used to measure optimism (Lyrakos et al., 2010).

- **Self-Description Questionnaire (SDQ)**

SDQ is a 6 point multidimensional questionnaire (Marsh et al., 1984) which measures 7 aspects of self-concept and includes 70 items. SDQ is a reliable tool which can differentiate the different aspects of self-concept (Marsh et al., 2004). The questionnaire has been validated in Greek population, showing a factorial structure similar to the original questionnaire (11-factors) (Tsorbatzoudis et al., 2005).

Procedure

At first, individuals read the informed consent. After that, participants completed the psychometric tools on a voluntary basis. The procedure lasted 30 minutes maximum. Three months after the initial assessment, 30 participants from the original sample were re-assessed in order to measure the test retest reliability of SDS scale in presence of the researcher.

3   Preliminary results

The Chi-square test was used to compare the percentages of the scales' sum scores (Table 2), whereas correlations were tested by Pearson r due to the normal distribution of the sample. The following psychometric properties were evaluated: internal reliability assessed by Cronbach's alpha coefficient (minimum acceptable value is 0.7) (Cronbach 1951; Clark and Watson 1995), concurrent validity (correlations be-

tween SDS and each one of the above scales) and test retest reliability. Statistical analysis was conducted by SPSS v 17.

Demographics

Our final sample consisted of 52 men and 182 women aged from 18-32 years. The majority of them were non married (59.6%), and they currently studying a university degree (70.2%). The demographic data of the sample is presented in Table 1, whereas mean scores and standard deviations of the scales' administered are placed in Table 2.

Internal consistency

All the scales were used had a satisfactory internal consistency, BDI (α= .86), DSS (α= .86), SDQ (α = .95) and RSES (α= .84). The SDS had very good internal consistency. Cronbach's alpha coefficient for the 12 items was high (α= .86), which means that each item is correlated positively with the sum of the other items of the scale (all r [1090]> .5) and significantly (all ps< .0001).

Construct validity

The SDS scores significantly correlated with the BDI (r = .631, p = .00), RSES (r =-.703, p < .00), LOT (r = -.416, p = .00) and SDQ (r = -.410, p = .00), although there was not a significant correlation with DSS (r = .004, p = .955).

4   Conclusions

The aim of the present study was to validate the SDS in the Greek population. The Greek version of SDS was found to have high internal reliability and satisfactory construct validity. Although more analyses need to be conducted to test the psychometric properties of the scale, the preliminary results suggest that this tool is both valid and reliable for self-disgust measurement in Greek population. One significant differentiation between English and Greek version of the SDS is that the Greek version did not correlate significantly with the DSS unlike the English version (Overton et al., 2008). In order to explain this finding, it can be assumed that Greek participants evaluated the way they feel about themselves differently to how they feel disgust about the scale's external disgust elicitors. Despite the fact that self-disgust constitutes a specific form of disgust, the participants of our study made a distinction between the feeling of disgust for external stimuli and the disgust for the self and behavior. Furthermore, taking into account that SDS Greek version was correlated with BDI, we can conclude that the feeling of disgust is more close to depressive symptomatology and the negative feelings for the self rather than the sense of disgust towards external elicitors.

As a conclusion, the Greek SDS version can be used in Greek population and presumed as a reliable index for self-disgust measure.

Table 1. Demographic variables for all study participants.

Variables	Sample (n=234)
Mean	22. 4
S.D.	3.3
Range (years)	18-32
Men % (n)	22.1.1% (52)
Women % (n)	77.4% (182)
Education (years)	13-20
Mean	14.7
S.D.	1.46
Range	13-24
Marital status % (n)	
Married	1.3% (3)
Non married	59.6% (140)
In relationship	38.7% (91)

Table 2. Scales' mean scores and SD

	Mean scores	Standard deviation	p value
SDS	28.00	10.59	p< .05
Beck II	9.92	7.75	p< .05
Disgust scale	62.97	16.28	p< .05
Rosenberg self esteem	21.0	4.88	p< .05
SDQ	271.94	54.49	p< .05
LOT- R	22.33	4.13	p< .05

References

Beck, AT (1967) *Depression: Causes and treatment.* University of Pennsylvania, Philadelphia

Beck, AT (1995) *Cognitive therapy: Basics and beyond.* Guilford Press, New York

Clark, LA, & Watson, D (1995) Constructing validity: Basic issues in objective scale development. *Psychological Assessment.* 7, 309–319

Fountoulakis, KN, Iacovides, A, Kleanthous, S, Samolis, S, Gougoulias, K, St Kaprinis, G, & Bech, P (2003) The Greek translation of the symptoms rating scale for depression and anxiety: preliminary results of the validation study. *BMC psychiatry.* 3, 1, 21

Fountoulakis, KN, Papadopoulou, M, Kleanthous, S, Papadopoulou, A, Bizeli, V, Nimatoudis, I, Iacovides, A, & Kaprinis, GS (2006) Reliability and psychometric properties of the Greek translation of the State-Trait Anxiety Inventory form Y: preliminary data. *Ann Gen Psychiatry.* 5, 2, 1-10

Haidt, J, McCauley, C, & Rozin, P (1994) Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences.* 16, 701-713

Koumi, I (1994) *Self-values and academic self-concept of Greek Secondary School pupils.* Ph.D thesis, University of Bristol

Koumi, I, & Tsiantis, J (2001) Smoking trends in adolescence: Report on a Greek school-based peer-led intervention aimed at prevention. *Health Promotion International.* 16, 65-72

Olatunji, BO, Ciesielski, BG, Wolitzky-Taylor, KB, Wentworth, BJ, & Viar, MA (2012) Effects of experienced disgust on habituation during repeated exposure to threat-relevant stimuli in blood-injection-injury phobia. *Behavior therapy.* 43,1, 132-141

Olatunji, BO, Haidt, J, McKay, D, David, B (2008) Core, animal reminder, and contamination disgust: Three kinds of disgust with distinct personality, behavioral, physiological, and clinical correlates. *Journal of Research in Personality.* 42, 1243-1259

Overton, PG, Markland, FE, Taggart, HS, Bagshaw, GL, & Simpson, J (2008) Self-disgust mediates the relationship between dysfunctional cognitions and depressive symptomatology. *Emotion.* 8, 3, 379-385

Phillips, ML, Senior, C, Fahy, T, & David, AS (1998) Disgust—the forgotten emotion of psychiatry. *Br J Psychiatry.* 172, 373–375

Powell, P.A, Overton, P.G. & Simpson (2014) *The Revolting Self: Perspectives on the psychological and clinical implications of self-directed disgust,* J. Karnac (in press)

Power, MJ, & Dalgleish, T (2008) *Cognition and Emotion: From Order to Disorder (2nd ed.),* Psychology Press, Hove

Robins, RW, Hendin, HM, & Trzesniewski, KH (2001) Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and social psychology bulletin*. 27, 2, 151-161

Rosenberg, M (1965) *Society and the adolescent child*, NJ: Princeton University Press, Princeton

Rozin, P, Haidt, J, & McCauley, CR (2008) Disgust, In Lewis, M, Haviland-Jones, JM, & Barrett, LF (eds), *Handbook of emotions, 3rd ed.*, Guilford Press, New York, 757-776

Rusch, N, Schulz, D, Valerius, G, Steil, R, Bohus, M, & Schmahl, C (2011) Disgust and implicit self-concept in women with borderline personality disorder and post-traumatic stress disorder. *European Archives of Psychiatry and Clinical Neuroscience*. 261, 369–376

Schienle, A, Schäfer, A, Stark, R, Walter, B, Franz, M, & Vaitl, D (2003) Disgust sensitivity in psychiatric disorders: A questionnaire study. *The Journal of Nervous and Mental Disease*. 191, 831–834

Schmitt, DP, & Allik, J (2005) Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: exploring the universal and culture-specific features of global self-esteem. *Journal of personality and social psychology*. 89, 4, 623

Sotiropoulou, V, Chalimourdas, T, Matsouka, E, Niarou, D, Kalogeraki, L, Antoniou, M, Mourikis, I, & Vaidakis, N *DS in Greek language. Reliability and Factor analysis*. 3rd Congress of Cognitive therapists, Thessaloniki, 2012

van Overveld, M, de Jong, PJ, Peters, ML, & Schouten, E (2011) The Disgust Scale-R: A valid and reliable index to investigate separate disgust domains?. *Personality and Individual Differences*. 51, 3, 325-330

Whelton, WJ, & Greenberg, LS (2005) Emotion in self-criticism. *Personality and individual differences*. 38, 7, 1583-1595

Appendix The Self-disgust Scale (SDS) in Greek

Αυτό το ερωτηματολόγιο αφορά στα συναισθήματα που έχεις για τον εαυτό σου. Αφού διαβάσεις τις προτάσεις, κύκλωσε τον αντίστοιχο αριθμό ανάλογα με τους ακόλουθους ορισμούς: 1 = Συμφωνώ απόλυτα; 2 = Συμφωνώ αρκετά; 3 = Συμφωνώ εν μέρει; 4 = Ούτε συμφωνώ ούτε διαφωνώ; 5 = Διαφωνώ εν μέρει; 6 = Διαφωνώ αρκετά; 7 = Διαφωνώ απόλυτα.

	Συμφωνώ απόλυτα						Διαφωνώ απόλυτα
1. Βρίσκω τον εαυτό μου αποκρουστικό.	1	2	3	4	5	6	7
2. Είμαι περήφανος/η για τον εαυτό μου.	1	2	3	4	5	6	7
3. Ο τρόπος που φέρομαι με κάνει να απεχθάνομαι τον εαυτό μου.	1	2	3	4	5	6	7
4. Μισώ αυτό που είμαι.	1	2	3	4	5	6	7
5. Απολαμβάνω την παρέα άλλων ανθρώπων.	1	2	3	4	5	6	7
6. Μου αρέσει η εμφάνιση μου.	1	2	3	4	5	6	7
7. Γενικά, οι άνθρωποι με αντιπαθούν.	1	2	3	4	5	6	7
8. Μου αρέσει να βρίσκομαι έξω.	1	2	3	4	5	6	7

	1	2	3	4	5	6	7
9. Αισθάνομαι καλά με τον τρόπο που συμπεριφέρομαι.	1	2	3	4	5	6	7
10. Δε θέλω να με βλέπουν.	1	2	3	4	5	6	7
11. Είμαι κοινωνικό άτομο.	1	2	3	4	5	6	7
12. Συχνά κάνω πράγματα που τα θεωρώ αηδιαστικά.	1	2	3	4	5	6	7
13. Μερικές φορές αισθάνομαι ευτυχισμένος/η.	1	2	3	4	5	6	7
14. Είμαι αισιόδοξο άτομο.	1	2	3	4	5	6	7
15. Με ενοχλεί να κοιτώ τον εαυτό μου.	1	2	3	4	5	6	7
16. Μερικές φορές αισθάνομαι λυπημένος/η.	1	2	3	4	5	6	7
17. Δεν μου αρέσουν κάποια χαρακτηριστικά της προσωπικότητάς μου.	1	2	3	4	5	6	7
18. Η συμπεριφορά μου απωθεί τους ανθρώπους.	1	2	3	4	5	6	7

The effect of processing demands on hostile attribution bias (HAB)

Luiza Shahbazyan[1]

[1] Department of Cognitive Science and Psychology, New Bulgarian University
`l.shahbazian@gmail.com`

Abstract. The hostile attribution bias (HAB), or the tendency to attribute hostile intentions to others' provoking, but ambiguous behaviour, has been recognized as an important proximal cause of aggression. According to the dominant social information processing account of HAB (Dodge, 2006), the bias results from the chronic access to hostile knowledge. However, some recent findings suggest (Rosset, 2008) that an increase of processing demands can induce HAB even in non-aggressive participants. The aim of this paper is to briefly review the evidence for the role of processing demands in hostile attribution and to outline the design of an experiment that might reveal this role. The results of the proposed experiment are expected to get to a better understanding of how people make hostile attributions and to reveal the role which the interaction between processing demands and knowledge accessibility plays in such attributions.

Keywords: aggression, hostile attribution bias, processing demands

Introduction

It is well known that aggressive behaviour has multiple determinants: genetic, physiological, emotional, cognitive, social and ecological among others. According to the social information processing accounts of aggression and violence, however, each particular act of aggression is preceded by a number of information processing steps, whose output serves as a proximal cause of the aggressive behaviour (Dodge, 2006; Heusmann, 1998). More specifically, a robust correlation (about .76) has been established between the tendency to attribute hostile intention to a person, even if the action does not indicate a clear motivation, and the likelihood to respond aggressively (see Dodge, 1980; Orobio de Castro, Veerman, Koops, Bosch, & Monshouwer, 2002, for a review). In a typical study, participants were presented with one or more ambiguous provoking situations and asked to point out if the negative outcome was an accident or a result of the hostile intention of the instigator. Aggressive children and adults from both clinical and non-clinical populations are more likely to make hostile intentional attributions in comparison to matched control sample (Epps & Kendall, 1995; Orobio de Castro et al., 2002). It was proposed that aggressive children exhibit

HAB as a result of persistent activation of chronically accessible hostile knowledge structures (Graham & Hudley, 1994; Burks, Laird, Dodge, Pettit, & Bates, 1999; Dodge, 2006). However, some recent findings suggest that intentional interpretations may be the default of how people in general reason about others' intentions under time constraints (Rosset, 2008). These findings reveal an important limitation of current models. None of these models can specify how processing demands influence hostile intentional attribution, although some authors suggested possible connection between the two (Dodge, 2006). The aim of this paper is to briefly review existing evidence suggesting a possible role of the processing demands and propose an experimental design to explore the effect of processing demands on HAB.

According to Crick and Dodge (1994), existing hostile schemes distort the attribution of intent in aggressive children by a process that is "rapid, automatic, irrational and probably classically conditioned" (p.79). In this light, it is truly surprising that almost the entire research on HAB is based on experimental paradigms involving reflection (Crick & Dodge, 1994; Fontaine & Dodge, 2006), while the phenomenon is considered to account for fast and effortless responses like yelling at someone who cut you off while driving (Matthews & Norris, 2002). For instance, in one of his paradigmatic studies, Dodge (1980) asked children to participate in puzzle assembling, who were randomly assigned to three control conditions: hostile, benign and ambiguous. The puzzle of each child was destructed by a unseen peer, who, through simulated "live" self-talk, was heard to be acting with a hostile intent, with a benign intent, or ambiguously. In the hostile intent condition all children reacted with aggression, in the benign condition all children restrained from aggression. In the condition where the intention of the peer remained ambiguous, aggressive and nonaggressive boys diverged in their reactions, so that aggressive boys reacted as if the peer had acted with hostile intent, while nonaggressive boys behaved as if the peer had acted benignly. Similar differences between aggressive and non-aggressive boys were obtained when children were asked to comment a peer's behaviour in hypothetical situations. In both paradigms no speeded judgment or immediate action was required from the participants, so it could be assumed that they engaged in a controlled, effortful processing of the situation. However, it was demonstrated that reflective and fast responses often differ and the latter may be closer to the actual behaviour (Rabiner, Lenhard, & Lochman, 1990). After all, provocative situations rarely invite reflective reasoning and they are often coupled with highly emotional response that is also known to increase information processing demands and induce automatic, preemptive processing (Lemerise & Arsenio, 2000).

While reflective responses may be very informing for the deficiency of aggressive participants in terms of explicit benign social knowledge (Rabiner et al, 1990), recent evidence outside the developmental field suggested that processing demands may be important determinant of whether intention will be attributed to an ambiguous action. Rosset (2008) asked college students to evaluate simple sentences containing actions. Some of these actions were clearly accidental ("She broke her leg"), some were clearly intentional ("She proofread the paper"), and some were ambiguous ("She broke the vase"). When adults need to make speeded judgments, they become much more prone to judge ambiguous actions as intentional. Rosset & Rottman (2014) went further to

suggest that people have initial impulses to interpret actions intentionally and their overcoming in order to perceive an action as incidental, is demanding.

However, Buon, Jacob, Loissel & Dupoux (2013) failed to replicate Rosset's (2008) findings. In their study, instead of short sentences, participants were presented with short animations where one of the characters hurts the other either on purpose or accidentally. While under difficult concurrent task participants tend to judge more harshly the actor despite that her actions were not intentional, under the low processing demand condition, processing demand did not produce more intentional attributions to the accidental action. One possible explanation of this is that a bias towards intentionality is observed only when sufficient ambiguity is present. Although Rosset (2008) reported the effect also with actions that are rated as unambiguously accidental, some of the sentences which she categorized as unambiguous can indeed be interpreted as ambiguous like "She broke her cell phone". The present ambiguity could explain the increase of intentionality under speeded judgment. To replicate Rosset's findings and extend it to test the effect of processing demands on stimuli that are typically used to assess HAB, the following experimental design is proposed.

Proposed Methodology

In order to test the hypothesis that a higher processing demand results in more intentional attributions, a Latin Square within subjects design is proposed. Young adults will be invited to participate in an experiment about memory retrieval. The independent variable will be processing load (high vs. low) and the dependent variable will be the interpretation of the actor's behaviour (intentional vs. unintentional). Additionally, participants will complete Buss and Perry's (1992) Aggression Questionnaire (AQ), translated and adapted in Bulgarian by Калчев (2005) to control for individual differences in aggression.

The situations to be evaluated will be adopted from Comb, Penn, Wicher, & Waldheter's (2007) Ambiguous Intentions Hostility Questionnaire (AIHQ). The questionnaire contains 15 situations that are either clearly intentional, or clearly accidental, or ambiguous in respect to the actor's behaviour. The situation descriptions will be translated into Bulgarian and will be given to university students to categorize them in order to avoid any possible overlap between the categories. Taking into account the results of Rosset (2008) and Buon et al (2013), it is predicted that processing demands will increase intentional attributions in ambiguous sentences, but not in the clearly accidental and in the clearly intentional ones.

The experiment will be run on a computer using E-Prime software, developed by Psychology Software Tools. It will consist of 8 practice trials and 12 experimental trials. Half of the trials will involve high processing demand by requiring the participant to remember four 2-digit numbers presented before the target story and report them after they entered the interpretation of the story's actor behaviour. The other half of the trials will involve remembering one 2-digit number, which corresponds to the low processing demand condition. Similar procedure was used successfully by Lin,

Keysar and Epley (2010) to diminish participants' ability to imagine themselves in someone else's perspective.

A pilot study will be done in order to establish the exposure time that will guarantee that the correct attributions of intention in the accidental and in the intentional stories will be significantly above the chance. At the end, the participants will be debriefed.

Conclusions

The suggested experiment will provide the first empirical test of the hypothesis that demanding processing can induce HAB. Although Dodge (2006) has supposed that executive control may play a role in mediating hostile attribution bias, his model does not make specific predictions about the mechanism of the supposed mediating effect. Fontaine & Dodge (2006) suggested that forcing aggressive children to slow down may decrease the tendency to make hostile attributions and forcing rapid responding may induce aggressive behavior in nonaggressive children. Although these suggestions are in line with the expected findings of this study, they refer to the process of decision making, which is conceptualized as taking place much later than the attribution of intent. In other words, rapid responding was suggested to influence the process of decision making when the child has already established that the situation is hostile, but the response is yet to be enacted. In contrast, the present study aims to explore how increasing the processing demands affects the very attribution of intention, mimicking conditions of fast and spontaneous real life attributions.

An important question still to be answered is why the bias towards intentionality is observed under processing demand. One possibility is that this is the primitive, automatic mode of how people understand intentions as suggested by Rosset (2008). Another possibility is that this bias is a result of retrieving the most accessible knowledge and intentional actions may be remembered better than unintentional ones. Indeed, Rosset (2008) reported preliminary findings that intentional sentences are remembered better than unintentional ones. Further research is needed to confirm this pattern and assess how processing demands and knowledge interact to produce the observed findings.

One important limitation of the proposed study is that, similarly to the criticized experiments, it relies on the explicit responses of the participants. More implicit measures may be employed to assess the divergence between low demand and high demand processing like eye tracking and priming.

The expected findings have important implications for the understanding of the ethology of HAB and for the design of successful interventions to remedy this bias. For instance, working memory training may alter HAB in children and adults more effectively than more popular approaches like cognitive restructuring and social skills training, which target more controlled aspects of cognition such as problem solving and decision making, but with which such a change is very difficult to achieve (Dodge, 2006).

References

Buon M., Jacob P., Loissel E., Dupoux E. (2013). A non-mentalistic cause-based heuristic in human social evaluations. *Cognition. 126*, 149-55.

Burks V., Laird R., Dodge K., Pettit G., & Bates J. (1999). Knowledge structures, social information processing, and children's aggressive behavior. Social Development , 8, 220-236.

Combs D. R., Penn D. L. , Wicher M., Waldheter E. (2007). The Ambiguous Intentions Hostility Questionnaire (AIHQ): a new measure for evaluating hostile social-cognitive biases in paranoia. *Cognitive Neuropsychiatry, 12*, 128-43.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115,* 74–101.

Dodge, K. (2006) Translational Science in Action: Hostile Attributional Style and the Development of Aggressive Behavior Problems. *Development and Psychopathology, 18*, 791-814.

Fontaine R, & Dodge K. (2006).Real-time decision making and aggressive behavior in youth: A heuristic model of response evaluation and decision (RED). *Aggressive Behavior, 32,* 604–624.

Graham S, & Hudley C. (1994) Attributions of aggressive and nonaggressive African-American male early adolescents: A study of construct accessibility. *Developmental Psychology, 30*, 365–373.

Huesmann, L. R. (1998). The role of social information processing and cognitive schema in the acquisition and maintenance of habitual aggressive behavior (pp. 73-109). In R. G. Geen & E. Donnerstein (Eds.), *Human Aggression: Theories, Research, and Implications for Policy.* New York: Academic Press.

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development, 71*, 107–118.

Lin S., Keysar B., & Epley N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology, 46*, 551–556.

Matthews. A. & Norris F.H. (2002) "When is believing "seeing"? Hostile attribution bias as a function of self-reported aggression", *Journal of Applied Social Psychology, 32*, 1-32.

Orobio de Castro, B., Veerman, J. W., Koops, W., Bosch, J. D., & Monshouwer, H. J. (2002). *Hostile attribution of intent and aggressive behavior: A meta-analysis. Child Development, 73*, 916-934.

Rabiner, D. L., Lenhart, L., & Lochman, J. E. (1990). Automatic vs. reflective social problem solving in relation to children's sociometric status. *Developmental Psychology, 26*, 1010-1026.

Rosset, E. (2008). It's no accident: Our bias for intentional explanations. *Cognition, 108,* 771-780.

Калчев П. (2005). Българска версия на въпросника на А. Бъс и М. Пери за оценка на агресията. *Психологични изследвания, 2*, 17-40.

Investigating executive control and attention in relation to workload in younger and older employees

Bujar Gallopeni[1], Prof. Rod Nicolson[2], and Dr. Antonia Ypsilanti[3]

[1]Research Student / Ph.D Candidate, South East European Research Centre (SEERC), Research Centre of the International Faculty of the University of Sheffield, CITY College, bgallopeni@seerc.org, bgallopeni1@sheffield.ac.uk

[2]Proffesor, Department of Psychology, University of Sheffield, r.nicolson@sheffield.ac.uk

[3]Lecturer, Department of Psychology, the University of Sheffield International Faculty, City College, aypsilanti@city.academic.gr

Abstract. This PhD research project aims to investigate how cognitive processes such as executive control and attention influence the employability of aging people. In addition, the research projects aims to investigate the influence of other psychological and job-related factors, such as personal motivation, job satisfaction and work readiness with their potential role as mediators/moderators to cognitive processes and work performance of aging people. The research further seeks to present comparative analysis between young and old study groups in order to meet the hypothetical expectations. The research methodology such as correlation and regression analysis, t-test and ANOVA examination methods are expected to derive research conclusions in relation to asked hypotheses.

Key words: *executive control, aging people, personal motivation, job satisfaction, job performance*

1. Background to research topic

The European Union is facing great challenges due to the economic crisis and the effect of the aging population worldwide. Organizations will be facing challenges and opportunities in the following decades that relate to the management of employees from different generations in many aspects (e.g., working together, knowledge exchange, innovation etc.). It is further anticipated that as a result of demographic changes organizations will continue to have the need of using employees longer in service, particularly when it comes to issues such as innovation, change and capacity

building (Silver Projects, accessed 30-03-2014)[1]. Consequently, many developed countries have been adopting retirement policy to favoring extension of work beyond retirement (D'Addio et al, 2010). This policy encompasses stimulus measures, such as pension raise (Chomik and Whitehouse, 2010), more benefits for late retirement and disfavoring early retirement (D'Addio et al, 2010), etc.

Not are only important policy measures to stimulate the elders' later employment, it is also crucial to examine the individual/cognitive factors and other work-related factors that contribute to a better work performance at old ages. It is thus very important to investigate, with a multidisciplinary approach, working and aging particularly factors that affect job performance in older and younger adults. These factors may vary from cognitive dimension, including working memory, attention, etc., to factors related to work itself, such as work motivation, satisfaction, readiness, and so on.

2. Literature review

Many studies have been focused in the last decades in examination of interplay among cognitive skills, such as executive functions, work performance and other work factors among older people in comparison to their younger counterparts. For example, Spitulnik (2006) gives an overview on cognitive development needs among older and younger people in relation to their performance. Although previous research points out that cognitive capacity declines with age, which makes elder people process information more slowly and have slower memory, older people perform better on problem-solving tasks (Warr, 2001), which is relied on the accumulated knowledge (Salthouse and Maurer, 1996). These examples imply how both fluid and crystalized intelligence develop through age and how they affect individual's performance at work. It particularly implies *the role of crystalized intelligence to work performance of elderly people.*

Several studies have documented that fluid intelligence is more affected by aging than crystallized intelligence (for a review see Ypsilanti & Vivas, 2012). Horn and Cattell (1966) explained that crystallized intelligence involves general acquired knowledge and vocabulary and is related to linguistic ability and expression. Fluid intelligence refers to ability to think logically and to solve novel problems, to reason and to draw conclusions. Results from studies investigating memory difficulties in younger and older healthy workers indicated controversial results, with most studies point to idea that if memory difficulties exist they are evident after the age of 70 years (e.g., Cabeza et al., 2004; Blacker et al., 2007). Speed of processing has also been extensively investigated in older workers indicating that there is evidence for deterioration in speed with aging. This project will further explore factors that influence job performance from a cognitive psychology perspective. Specifically, it

---

[1] http://www.intergenerationallearning.eu/project-silver/

aims to investigate executive function and attention in relation to workload in younger employees using cognitive models. Since a large body of research highlights the importance of cognitive training through lifelong learning in successful ageing and increased job-related performance (Kemper, 1994; Lawrence, 1996; Schaie, 1994) it is vital to understand more about the mechanisms that support these functions. More specifically, this thesis will explore **executive control**, **attention switching** (selective attention) i.e. task management (e.g., Dismukes & Nowiski, 2007) using cognitive models in relation to workload in younger and older employees.

Executive control is described as the ability to guide thought and action in accordance with internal goals (http://decisionneurosciencelab.org/the-dnl-lab/executive-control-and-the-prefrontal-cortex/, accessed: 06-04-2014). It has been considered as a multidimensional construct, having four components, such as volition, planning, purposive action and effective performance (Lezak, 1995; Spreen & Strauss, 1998), which often involve attention and working memory, cognitive flexibility, decision making, judgment, and behavioral regulation.

The Grundtvig SILVER project recently outlines three types of learning that elder people in organizations should focus on: learning to learn, learning to innovate, and learning to build knowledge, as key competences to increase their employability (http://www.intergenerationallearning.eu/introduction-to-the-toolbox/). In this regard important role play also the personal motivation and the ability to learn (De Grip et al., 2004; Van der Heijde and Van der Heijden, 2006). Kanfer and Ackerman (2004) suggest that there is no decline of work motivation with aging, and that work motivation concept develops on the same principles among midlife and older people as to the younger people. They further suggest that older people are particularly looking to personal allocations of efforts on work which promotes particularly their self-efficacy and self-concept. *These examples imply the role of intrinsic motivation in increasing work readiness among older people.*

Consequently, this project seeks to examine the relations between intrinsic motivation, crystalized intelligence as well as their (mediating/moderating) contribution to improve executive functions towards work performance (H1).
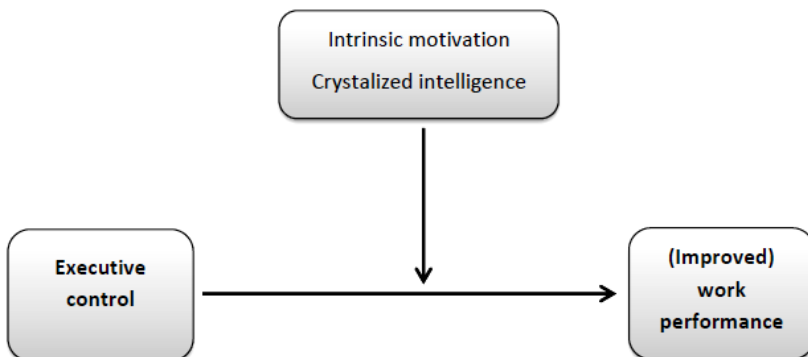
Fig. 1. Relation of executive control to work performance among old adults influenced by possible facilitating factors

Recent studies show that executive control performance improved when receiving feedback (Drueke et al., 2012). ***In this regard, (positive) feedback, as a performance improvement stimulus, might increase personal motivation which in turn might increase the level of job satisfaction and work readiness, and all this chain of influence might be serving as catalyst to tune up the executive functions among elder employees and thus increasing their employability (H2).***
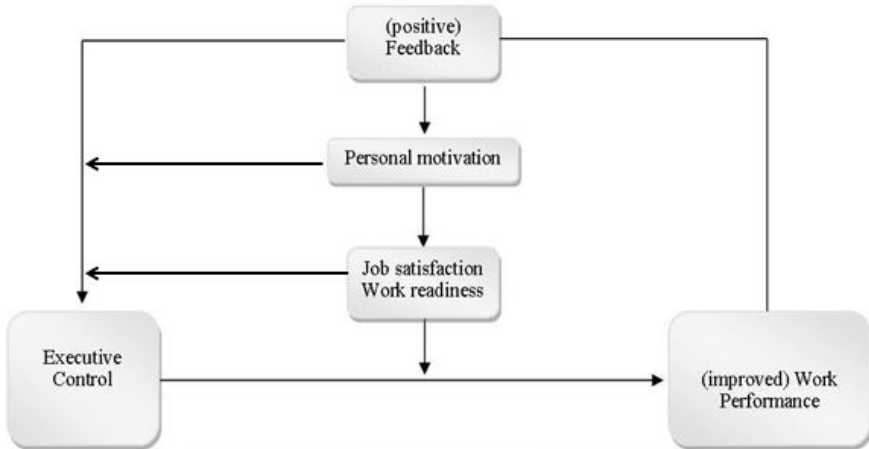
Fig. 2. Executive control in relation to improved work performance as moderated/mediated by positive feedback/personal motivation/job satisfaction/ work readiness.

3. Proposed Methodology

3.1 Sample

Two group cohorts of participants are proposed to be included in the study, *baby boomers* and *generation x*. Around two hundred (200) participants are proposed to be included in the study in order to be able to undertake needed analyses. Participants might be coming from different organizational types, such as public institutions (governmental, health, education, etc), private sector (production, consultancy, trading, etc) via different occupations (managerial, technical, policy making, etc) in order to reflect differences in the independent variables such as fluid/crystalized intelligence among age groups. The gender aspect will be also included in order to test potential differences among this dimension as well.

3.2 Materials and data collection

A combination of standardized cognitive tests and self-report questionnaires are proposed to collect data from participants. Cognitive tests will be used mainly to measure the executive control and the crystalized intelligence variables, while the self-report questionnaires and tests will be used to measure personal motivation, job satisfaction, work readiness, and so on.

3.3 Proposed procedure and analysis

Between group analysis using ANOVA or/and T-test is proposed as methodology to reflect to the potential differences examining levels of executive control and attention switching to younger and older employees. Differences are to be observed in relation to variables such as age (baby boomers, generation x), crystalized intelligence, and cognitive processes (executive controlling, attention/task switching).

Attention switching variable could be measured using Attention Switching Task (AST) test, which is a sensitive measure of frontal lobe and 'executive' dysfunction, and has 7 outcome measures, each of which can have various options applied to it. The AST measures cover latency, correct and incorrect responses, commission errors, omission errors, switch cost and congruency cost (http://www.cognitiveatlas.org/task/Attention_Switching_Task, accessed: 07-04-2014). This test will not only show the capability of aging people in performing level in switching tasks at work, but it might also give a good comparison to younger adults as well.

4.   Conclusions

Recent demographics show that the world population, particularly in developed countries, is getting older, and that countries are considering policy review for optimizing the retirement of elderly to meet the labor market demands. The interest of organizations has started to increase in keeping older people further at work.

Previous research shows that older people are better performing in crystalized intelligence and that their (intrinsic) motivation is present when it comes to promotion of self-efficacy and self-concept. It is therefore of great importance that this research project examines the interplay between intrinsic motivation, crystalized intelligence and work performance, via influenced executive functions. Research has further shown that feedback at work plays a positive role onto improvement of performance of executive functions. It is on further interest of this project to examine the work/performance feedback influencing in turn the work motivation and job satisfaction which are assumed to play a mediating/moderating role in facilitating influence of improved executive functions towards work performance.

Adequate research methodology is proposed to collect data and derive analysis to examine the research assumptions.

5. References

Blacker, D. et al. (2007). Neuropsychological measures in normal individuals that predict subsequent cognitive decline. *Archives of neurology*, Vol. 64, No 6, p. 862-871.

Cabeza, R. et al. (2004). Task-independent and task-specific age effects on brain activity during working memory, visual attention and episodic retrieval. *Cerebral cortex*, Vol. 14, No 4, p. 364-375.

Chomik, R., and Whitehouse, E. R. (2010), 'Trends in Pension Eligibility Ages and Life Expectancy, 1950–2050', *OECD Social, Employment and Migration Working Papers,* No. 105, OECD Publishing, Paris, Organization for Economic Cooperation and Development.

Cremer, R. (eds). Work and aging: a European perspective. *Bristol: Taylor and Francis*, p. 31-46.

D'Addio, A. C., Keese, M., Whitehouse, E. (2010). Population ageing and labour markets. *Oxford Review of Economic Policy*, Vol. 26, Number 4, pp. 613–635

De Grip, A. et al.(2004). The industry employability index: taking account of supply and demand characteristics. International labor review, Vol. 143, No 3, p. 211-233.

Drueke et al. (2012). Can executive control be influenced by performance feedback? Two experimental studies with younger and older adults. *Frontiers in Human Neuroscience*. Vol 6, No 90.

Horn, J.L.; Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of educational psychology*, Vol. 57, No 5, p. 253-270.

Kanfer, R., Ackerman, P. (2004). Aging, Adult Development, and Work Motivation. *Academy of Management Review*, Vol. 29, No. 3, 440–458.

Kemper, H.C.G. (1994). Physical work and the physiological consequences for the aging worker. In: Snel, J.;

Lawrence, B.S. (1996). Organizational age norms: why is it so hard to know one when you see one? *Gerontologist*, Vol. 36, No 2, p. 209-220.

Lezak, M.D. (1995). Neuropsychological assessment. New York: Oxford University Press.

Schaie, K.W. (1994). The course of adult intellectual development. *American psychologist*, Vol. 49, No 4, p. 304-313.

Spitulnik, J. J., (2006), Cognitive Development Needs and Performance in an Aging Workforce. *Organization Development Journal*, Volume 24, Number 3.

Spreen, O., & Strauss, E. (1998). A compendium of neuropsychological tests: Administration, norms and commentary. New York: Oxford University.

Van der Heijde, C.M.; Van der Heijden, B. (2006). A competence-based multi-dimensional operationalization and measurement of employability. *Human resource management*, Vol. 45, Issue 3, p. 449-476.

Ypsilanti, A.; Vivas, A. (2011). Cognitive ageing in older workers and its impact on lifelong learning. In: Cedefop (ed.). Working and ageing: guidance and counselling for mature learners. Luxembourg: Publications Office, pp. 90-104. http://www.cedefop.europa.eu/EN/Files/3062_en.pdf [accessed 29.03.2014].

Executive functions and reading comprehension performance in children

Fatbardha Qehaja-Osmani[1] , Dr. Elisavet Chrysochoou[2], and Dr. Daniel J. Carroll[3]

[1] PhD student, South East European Research Centre (SEERC), University of Sheffield
fosmani@seerc.org

[2] SEERC Research Track 3 Leader, Lecturer – Psychology Department, University of Sheffield International Faculty, CITY College
echrysochoou@city.academic.gr

[3] Lecturer and Director of Learning and Teaching for the Department of Psychology, University of Sheffield, UK
d.carroll@sheffield.ac.uk

Abstract. Reading comprehension (RC) is a complex, dynamic process by which readers interact with text to construct meaning. It relies on lower level skills such as word decoding, fluency reading and vocabulary skills and also on higher level-skills such as inference making, literal comprehension and comprehension monitoring. Many research studies have found relationship between reading comprehension and executive functions (Sesma et al, 2009, Cain & Oakhill, 2007, etc). Executive functions (EF) are a highly complex, interdependent group of abilities that are involved in governing an individual's goal-directed behavior (Carroll et al, 2007). The present study aims at shedding more light on this area, by examining the relative contribution of four executive functions, inhibition, updating, cognitive flexibility and working memory to reading comprehension in children).

Keywords: *executive functions, children, reading comprehension, higher level skills, lower level skills.*

1   Introduction

One of the major concerns of policy makers, educators, and parents nowadays is adequate conditions that must be in place for all children to acquire quality knowledge and skills through education. Efficient language comprehension skills (oral or reading) provide the means for acquiring knowledge, also determining its consolidation and effective long-term use. Comprehension also allows us to be aware of and experi-

ence other (often fictional) worlds, to communicate successfully, and to achieve academic success (Cain & Oakhill, 2007).

Comprehension refers to the process of formation of a meaning –based representation of the text, often called a mental model or a situation model of the text (Gernsbacher, 1990; Kintsch, 1998; Zwann & Radvansky, 1998). Such models are seen as mental representations in the narrated situation rather than outside of it. A situation model is actually the representation of a text's "inner world" including information about agents, objects, events and other elements that help the reader or listener to interpret the given content, create new ideas or knowledge, through hypotheses and conclusions reached on the basis of the information provided.

Reading comprehension is not a passive; rather, it is the result of considerable cognitive activity. In the context of an 'construction –integration' (CI) model of text comprehension, Kintsch (1998) describes two major stages of text processing: the active development of a text base (construction phase), followed by the integration of the construction phase products in the text's mental representation (integration phase). Specifically, during the construction phase, there is a quick activation of linguistic, semantic, and world knowledge. The introduced knowledge is triggered by and related to the information explicitly stated in the currently processed text segment, with the help of the available working memory resources. In the context of the integration phase, knowledge units, carried over from the previous cycle, are linked with the products of the current construction phase products, in providing a complete, integrated mental model of the text (see also Eapson, Goldberg, Young, Geist, & Cutting, 2012).

Obviously, comprehension is among the most complex skills that children are called upon to master, coordinating various perceptual, linguistic, and cognitive skills (Braze, & Tabor, Shankweiler, & Mencl, 2007). In the following sections, processes that occur higher or lower in language comprehension will be described, along with evidence relating them to the cognitive resources of working memory and the executive system in the critical early years of development. Specifically, reference will be made to (a) lower level processes, such as word reading accuracy and speed, and semantic skills (e.g. vocabulary), as well as to (b) higher level processes, engaging the reader in the accurate storage information explicitly stated in the text (literal comprehension) and its integration to the text representation, in the generation of different types of inferences, the comprehension of similes, or in comprehension monitoring and control processes.

2   Lower level comprehension skills

Lower level skills refer to the abilities of children to semantically process words, relying on their vocabulary knowledge, and to decode or recognize them accurately and promptly. In some cases, children with comprehension difficulties seem to suffer in word-level processing while reading, which in turn exhausts the limited working memory resources (responsible for the temporary maintenance and processing of information) that are vital for higher-level processing. Moreover, if product of lower-

level processing is meaningless or if it is inaccurate or incomplete, inadequate input is provided to the higher level processing, affecting its efficiency (Kandeou, van den Broek, Helder & Karlsson, 2014). In the next section the contribution of each lower level skill will be discussed.

2.1 Vocabulary knowledge

Vocabulary knowledge and reading comprehension ability are highly correlated not just in children, but in adults as well (Carroll, 1993). Vocabulary significantly develops in the early years (e.g. Oullette, 2006; Cain, Lemmon, & Oakhill, 2004) in the context of both children's oral interactions with family and friends, and official, reading instruction at school. Children who aren't very good with words start out at a disadvantage, not only in speaking and listening skills, but also in learning how to read (Snowling, 2004). In the early stages of learning how to read, children start with focusing attention in phonics and later, they rely increasingly on word meanings to gain fluency in reading. Stothrad, Snowling, and Bishop (1998) followed the early reading development of 90 children between the ages of 4 years and 9 months and 6 years and 9 months. The researchers found that vocabulary knowledge was predicted by children's phoneme awareness, letter knowledge, and word recognition accuracy and speed. In the same study, vocabulary at age five did not seem to predict word recognition at age six; however, when the outcome was reading comprehension at age six, word recognition and vocabulary at age five were important predictors.

In a longitudinal study with English speaking children, Cain, Oakhill, and Bryant (2004) also found significant contributions of vocabulary knowledge to the prediction of reading comprehension variance at 8, 9 and 11 years. Similar findings are reported in a study with French speaking children conducted by Seigneur, Ehrlich, Oakhill, and Yuill (2000). We also know that vocabulary was a strong predictor of both oral and reading comprehension performance in Greek speaking children at the preschool and the elementary school phases (Chrysochoou & Bablekou, 2011; Chrysochoou, Bablekou & Tsigilis, 2011).

It should be noted that, in his research examining children from third grade through seventh grade, Stanovich (2000) found a reciprocal causal relation between reading and vocabulary: vocabulary growth leads to improved reading comprehension and amount of reading leads to vocabulary growth. The Matthew effect model of Stanovich (1986) provides a theoretical framework in which the development of individual differences in reading ability can be described and explained. The Matthew effect refers to the phenomenon that, over time, better readers get even better, and poorer readers become relatively poorer. This outcome has reference not only to word recognition and comprehension skills, but also to the development of cognitive skills related to reading, such as vocabulary and other (meta) linguistic skills. Students with greater vocabulary knowledge seem to have more opportunities to learn new words, through reading, inferring their meanings from context, as compared to their peers who possess lower vocabulary knowledge (e.g. Shefelbine, 1990; Senekal, Thomas, & Monker, 1995).

2.2   Word reading accuracy and speed

Given the various skills and processes involved in reading comprehension, poor vocabulary knowledge might only partially explain comprehension problems (see Pany, Jenkins, & Schreck 1982; Perfetti, 1985). During the comprehension of text children must also recognize the words fluently so as to promptly access their meaning (Cain & Oakhill, 2008). Word recognition can be accomplished by two strategies: (a) word decoding, namely via applying knowledge of grapheme-phoneme conversion, as well as knowledge of letter patterns, in order to appropriately pronounce written words, and (b) via sight-word reading, allowing the reader to recognize and pronounce written words correctly via addressing their orthographic representation. Decoding is the basic strategy applied when readers are presented with a transparent orthography (e.g. Greek, Italian, Albanian), whereas, sight-reading is a marker of proficient reading, as well as a valuable strategy in learning how to read in a language with an opaque orthography, such as English (e.g. decoding would not help a child to read the word vegetable). It should be noted, that in languages with transparent orthographies, children master decoding early on (within the first year of reading instruction); thus, reading fluency, taking reading speed, besides reading accuracy into account, seems to be a more appropriate word reading skill measure. Fluency or the ability to read words quickly and accurately is critical for reading comprehension, both for younger children, who still need to save resources for higher-order comprehension processes, which are still under development, and for older children, who are constantly required to use reading as a means for acquiring new information and knowledge.

A significant number of research studies have investigated the contribution of word decoding in reading comprehension (Krashen 2009; Keenan, Betjemann & Olson, 2008). Many research studies showed that vocabulary development plays an important role in word decoding. It has been proposed that the association between vocabulary and decoding is due to the role of vocabulary growth in the development of phoneme awareness (Goswami, 2001; Metsala, 1999; Walley, Metsala, & Garlock, 2003). In a study conducted by Sesma, Mahone, Levine, Eason & Cutting (2009) single word reading was a significant predictor of reading comprehension in the ages of 9 and 15.

There has been much recent attention to the possibility that the ease of reading acquisition may vary between languages, because some language have a more transparent orthography (e.g. Spanish, German etc.) while the others have an opaque orthography (e.g. English, Danish, Portugese). Wimmer and Goswami (1994) compared reading of digits, number names and non-words formed by exchanging the onsets and rimes of number names by 7-, 8- and 9-year-old children in German and English. Nonword reading was significantly slower and more error prone in English at all three age levels.

In an attempt to clarify the contribution of reading fluency to the reading comprehension performance of fourth graders, Jenkins, Fuchs, Van den Broek, Espin, and Deno (2003) showed that contextual word reading speed uniquely predicted reading comprehension whereas isolated word speeding did not. In a similar line, Cutting,

Materek, Cole, Levine & Mahone (2009), assessed children aged nine to 14 with two fluency measures of word reading efficiency and contextual word fluency. The children in this study were placed into one of three groups (typically developing group, those with general reading deficits, and the third group was composed of children with specific reading comprehension deficits) based on word reading and reading comprehension scores obtained during the screening visit. The group of GRD contained 18 participants, 17 participants were categorized in the group of S-RCD and 21 participants fit in the group of typically developing children. The researchers found isolated word fluency to be intact in most participants with specific reading comprehension deficits (thus, not accounting for their poor reading comprehension). With regard to fluency during text reading, children with specific reading comprehension indeed differed from the typically developed group. It is noted that the term specific reading comprehension difficulties refers to those children that struggle to understand what they read, despite having adequate word reading and vocabulary knowledge (see Oakhill & Garnham, 1988; Sesma et al. 2009; Torgessen, 2000). It has been suggested that, approximately 10% of young readers acquire age-appropriate word reading skills, but do not develop commensurate reading comprehension ability (Cain, 2004). For these children, comprehension difficulties often seem to emerge from inadequate processing of text information at a higher level (e.g. inference generation, monitoring of the comprehension pro, evaluation of its products, etc.).

3   Higher level comprehension skills

To date research studies have shown that besides mastering the lower level skills, successful comprehension significantly relies on higher level skills such as inference generation, simile comprehension, and monitoring - control of the comprehension process and products (Yuill & Oakhill, 1991; Oakhill & Garnham, 1988; Cain, Oakhill, & Lemmon,2004; Carretti, Borella, Cornoldi, & DeBeni, 2009). A child might, for example, be fully capable of converting graphemes to phonemes, thus reading fluently, as well as of accessing the meaning of each word in the mental lexicon rapidly, but at the same time fail to make sense of a text it is reading. There is evidence showing that comprehension difficulties can still manifest, even when lower level skills are well developed (Barnes, Huber, Johnston, & Dennis, 2007; Oakhill, 1993). Higher-level skills are more effortful because the reader has to integrate different parts of the text in order to develop an accurate and complete relevant representation (integration or explicit inference generation), as well as to link information from the text with relevant knowledge and experience, activated from his/ her long-term memory (e.g. in the case of implicit inference generation and simile comprehension; see Yuill & Oakhill, 1991). The next section higher-level comprehension processes, highlighting their role in reading comprehension.

3.1   Literal comprehension

An aspect of comprehension that has been widely measured in relevant studies is literal comprehension, namely the individual's   ability to recall information explicitly

stated in the story, not necessitating further, higher-order processing (e.g. inference generation). Such questions are assumed to set less demands, at least in terms of processing and recall, as compared to the above mentioned higher level skills (i.e. inference generation, simile comprehension, and comprehension monitoring and control). However, they reflect the quality of the construction and integration comprehension phases' products and more specifically, the efficient integration of the item to be recalled in the mental model constructed for the text. Literal comprehension or the effective representation of the basic text meaning is also crucial for both the on-line and the off-line processing at a higher level (e.g. respectively, drawing inferences while reading the relevant text segments or reading them, possibly also in relation to the following text segments or to relevant questions following the text read). Indicating the value of this comprehension skill, Perffeti, Landi & Oakhill (2005) argue that before one can conclude that inference making is a cause of poor comprehension, assurance is needed that the poor comprehender has constructed and maintained an effective representation of the literal meaning of the text. This means that many measures of text comprehension impose demands on memory, consequently poor comprhenders are considered to be participants who are likely to have poor memory for the text.

The literal meaning of the text is based on the short –term memory, considering that readers have to recall the sentence they have just read in order to further integrate it with other information. Poor comprehenders refer also to readers who find it hard to recall the immediately proceeding text and as such it appears difficult for them to link the current word or sentence with previous information.

On the other hand literal comprehension also relies on long term memory as readers will have to use the information in long –term memory to help them understand the text. Many studies (see Perffeti & Lesgold, 1977; Perfetti & Goldman, 1976) found that when examining children from third and fifth grades with a digit span task, language specific memory store which can hold information for rather longer period of time that the short term memory was an important component of literal comprehension. In line with these studies, Chrysochoou, Bablekou & Tsigilis (2010) examined 92 children of 7 and 9 years of age by using stories and questions tapping children's recall of information stated in the story. They found that children who were less skilled in reading comprehension were worse than the skilled once in the concurrent storage and processing of the text information. Inference making and comprheension monitoring are related to independent measures of working memory, suggesting that at least some of the difficulties experienced by poor comprehenders on these tasks may be due to working memory limitations (Cain, 2006). To illustrate this, poor comprehenders are particularly poor at spotting inconsistencies in text when several lines of text separate the two contradictory sentences. In these cases, the reader will only notice that something is wrong if he or she tries to integrate the just-read information with the existing situation model, rather than with the previous sentence (see Cain, 2009). In support of these findings, Barnes and Dennis (1996), found that literal memory for the text was a significant predictor for the coherence inference of the text in children 6 to 15 years old. However, this study has been contradicted by the finding of the study conducted by Oakhill, Barnes & Bryan (2001) in which they found that

literal memory of the text made no significant contribution for coherence inference generation.

Nevertheless, further work is needed to investigate whether literal comprehension is a predictor of other higher level skills such as inference making, comprehension monitoring and control in different age groups in children.

3.2.   Inferences

Research on inference generation dates back in 1970s, triggered by the development of Kintsch's the theory of text comprehension. During this period of research, inference generation was studied in relation to the goals and strategies of the reader, as well as relevant context effects. It was the second generation of research in inference making in 1980s (e.g. Fletecher & Bloom, 1988; van de Broek, 1990) that shed light on what readers do to draw an inference on-line, as they proceed through a text, with attention paid for example, to the demands set on the working memory during the inference generation process. In the 1990s inference making is further studied in relation to the integration of online and offline aspects of reading (e.g., Goldman & Varma, 1995; van de Broek, 1996). Specifically, online processing in reading is assumed to begin with a question set or an inference generation need identified, whereas the offline processing starts with finding answers to those questions while reading the text or after reading it.

During the reading process, simply identifying the meaning of individual words and sentences is insufficient. In normal reading situations, skilled readers construct text representations that are useful and coherent, that are easily accessed for future reference, and can be efficiently applied in a variety of situations (Rapp et al, 2007). In doing so, the reader is often required to go beyond the information explicitly presented in the text and produce new information; it is then that an inference is generated (St.George, Mannes & Hoffman, 1997). Specifically, in order to fully grasp the text's meaning, and construct its mental model, the comprehender has to find out how different pieces of information in the text fit together to achieve this goal. Often in doing so, one needs to rely on general knowledge and experience, stored in his/ her long-term memory. In the latter case, the inferences are referred to as implicit ones, in contrast to explicit inferences, requiring the integration of only explicitly stated text information (Yuill & Oakhill, 1991). Such higher-order processing involves the individual in identifying characters and their motives, in following the plot of the story, as well as in understanding information that is explicitly and implicit stated or simply implied in the story (Florit, Roch, & Levorato, 2011; Graesser, Singer, & Trabasso, 1994; Oakhill & Cain, 2007; van den Broek et al., 2005).

The development of inference making skills commences at a young age. Inference making is a part of children's everyday experiences, a central tool of making sense of the world. The generation of inference changes as the child grows, both in quality and quantity. For example younger children can generate inferences by connecting isolated events and facts in a story, whereas older children are more capable of connecting whole episodes and sections in the text (Cain & Oakhill, 1999). Furthermore, younger readers are less likely to generate inferences as compared to older children, even

when such higher-order processing is necessary (e.g. Catel & Simpson, 1999, Warren Treassur, 1978).

Cain, Oakhill, Barnes, & Bryant (2001) found a strong relation between comprehension skill and inference – making ability in 7-8 year-old children, even when knowledge was equally available to all participants. It seems that the difficulties some children face with inference generation is not due to the lack of relevant knowledge or poor memory for the text. The authors suggest that poor comprehenders seem to be able to integrate information at a local text level, but they are unable to interrelate the appropriate information in producing a coherent integrated model of the text as a whole. In addition, the authors demonstrated that less skilled comprehenders' difficulties with inference making are not just restricted to reading situations, but are also apparent in listening comprehension tasks; they thus seem to stem from higher order, rather than lower order, word-level processing.

Different types of inferences have been identified in the literature. Generation of coherence or necessary inferences, for example, requires the reader to integrate different pieces of information within the text. Skilled readers are characterized with the ability of making this type of inferences. Since coherence or necessary inferences are needed for a proper understanding of text, they must typically be made as texts are read (i.e. during the on-line process).For example, when readers are asked to think aloud while reading the text, they often give responses that explicitly reflect attempts to make coherences (Kandeou & van den Broek, 2005). Studies have revealed that in cases where causal information in the text are not explicitly stated, participants spend more time to read the text or sentence in order to build coherence (see Bloom, Fletcher, van den Broek,Reitz, & Shapiro, 1990). In such cases, where important information are not available in the current text, the participant will have to rely on previous knowledge to make sense of the text which consequently influences the online reading process.

Elaborative inferences produce information that is not that necessary for comprehension of the text's gist, but enrich its mental representation. The generation of such inferences can contribute to the development of a personal, emotionally driven relationship with the text, also facilitating its storage and recall (see Oakhill & Garnham; Block, Rodgers, & Johnson, 2004). Different from coherence inferences, elaborative inferences are considered to take place during the off-line processing. However, it should be noted that an inference cannot be classified as necessary or elaborative from its content or form but rather from the discourse context (see Oakhill & Garnham, 1988, pg. 23.). Many elaborative inferences can be made even from a single sentence text. Elaborative inferences are left to be drawn only if they are at some point possibly after having read the text – considered to be necessary. They are mostly generated with questions after the text has been read or when the reader needs to rely on the inference so as to make sense of other information provided in the text. As an off-line product of reading, elaborative inferences imply information that had to be retained in memory and become active when necessary. As such, future studies could explore differences in the measures of WM and EF contribution to both types of inferences, coherence and elaborative.

Cain, Oakhill, Barnes and Bryant (2001), examined children aged 7 – to- 8 and found that even when word reading has become fairly fluent and text comprehension has significantly developed, children are still learning to distinguish necessary from elaborative inferences.

Various factors affect the extent to which readers can allocate resources to the generation of inferences, regardless of their reading goals. Poor syntactic skills and limited working memory capacity (Just & Carpenter, 1992) may inhibit readers' ability to generate inferences (Oakhill, 1994). Specifically, working memory is considered to be the work space of inference making. It has been shown, for example, that making bridging inferences, to integrate information that is presented close together in the text is easier for typically developing children, compared to making bridging inferences across larger chunks of text (e.g. sentences or paragraphs), since the latter set greater demands on working memory (Johnston, et al., 2008). Actually, studies of children with age appropriate decoding skills, but poor comprehension performance, have found that making inferences across large textual distances affect poor comprehenders to a greater extent as compared to typically developing children (Cain, Oakhill, & Elbro, 2003; Cain, Oakhill, & Lemmon, 2004).

3.3 Comprehension monitoring and control

Skilled comprehenders may generate more inferences than less skilled comprehenders do, also possibly because they seem to be regularly and more efficiently monitoring the comprehension process, thus appropriately identifying the need to make inferences in order to fill in missing details (Cain, 2009; Cain et al, 2004). The ability to monitor one's level of understanding while reading is an important aspect of metacognition and a major component of learning. Measures of comprehension monitoring usually assess a reader's ability to detect inconsistencies in text, such as scrambled sentences, contradictory sentences, or statements that conflict with external information (world knowledge; see Cain, Brian & Oakhill, 2004). This kind of mental pulse-taking is important because it is a measure of progress towards a reading goal and a signal for comprehension failures (Paris & Meyers, 1981). Some researchers (e.g. Paris & Meyers, 1981) found that poor readers seem to be less aware of the existence and value techniques for regulating comprehension as compared to the skilled readers. Moreover, preschoolers seem to have difficulty in monitoring their own understanding while listening to stories or trying to understand and remember information while studying (Brown, 1978; Markman, 1979; Paris, 1978).

It has been suggested that monitoring of comprehension could be of two types: either be spontaneous, with the reader becomes aware of the need to evaluate his/her own understanding of the text, or directed monitoring, in which children are required to monitor the process. In an attempt to examine spontaneous comprehension monitoring, Paris and Meyers (1981) examined 42 fourth graders by asking them to read orally two stories that contained nonsense words and phrases and recorded their spontaneous corrections. In both cases the study showed that poor readers did not evaluate anomalous information as efficiently as good readers, even though they did engage in monitoring other text segments within the story they read that the same poor readers

also demonstrated less accurate comprehension and recall of the stories than good readers.

Interesting data with regard to the differences between poor and good readers in monitoring comprehension was also provided by a study conducted by Hacker (1997). Three groups of children were examined: eleventh grade poor readers, ninth grade poor readers and seventh grade good readers. When students were given the chance to find errors in texts they were reading, with the examiner pointing to the line containing an error (thus, appropriately directing the readers), students' performance improved. However even in that case, poor readers failed to improve as much as more skilled groups. These findings suggest that in low skilled monitoring readers relevant knowledge is not always used as needed. Since performance on comprehension monitoring tasks improves with age (Markman, 1981), looking at different ages of children (e.g. preschoolers and elementary school children of different ages would provide an important trajectory of children's developing skills in comprehension.

Comprehension control is an off-line process, relying on one's metacognitive awareness, as it regards the evaluation of the comprehension products off-line, when the individual is directed to. Comprehension control questions provide children with information that was inconsistent with the gist, thus evaluating text representation adequacy and metacognitive control processes. These provocative inaccurate questions can tap children's ability to control the meaning of the text they just read.

4.   Executive functions

Many studies (e.g. Cutting, Matterek, Cole, Levine & Mahone, 2008) have shown that executive functions, also called EFs, are associated with reading comprehension. Executive functions (EF) are control processes involved in the regulation of cognition (Miyake et al., 2000). Executive functions include working memory, switching or mental flexibility, and inhibition. Indeed, weaknesses in executive functions may seriously impair the reader's ability to perform the cognitive processes, important for adequate comprehension. Executive function is recognized as a critical component of children's cognitive and social development (Carlson, Davis & Leach, 2005). The first 5 years of life play a critical role in the development of executive functions (EFs). During this period, core components of EF develop, forming a critical foundation that will set the stage for the development of higher cognitive processes well into adulthood (Garon, Bryson, & Smith, 2008). Executive functions seem to improve sequentially through childhood (Anderson et al , 2001).

The current study will be based on the model of Miyakee and colleagues (2000) on executive function, as described above. This theoretical framework, concludes that the three core EF functions, although clearly distinguishable, they do share some commonalities. Thus the authors suggest that executive functions are separable but moderately correlated construct, therefore suggesting both unitary and non-unitary components of this system. Miyake et al. (2000) used confirmatory factor analysis (CFA), a structural equation modeling technique, to assess the validity of their model. The benefit of the CFA is that it is theoretically reliable, so that researchers can explicitly test their model against competing models. Miyake et al. argued that part of the diffi-

culty in studying components of EF is that the measures are not pure. Thus, according to this model, using different measures of the same EF component is assumed to be a purer measure of EF construct. Miyake et al found that the reliability of the data will be higher if each of the EF components is measured by three common EF measures .

In the next sections I describe the development of each executive function separately. Furthermore, in each EFs section I discuss its contribution to language comprehension in childhood.

4.1.   Cognitive Inhibition

Cognitive inhibition, also referred to as inhibitory control is a mechanism that allows children to stop the activation of the current task. Cognitive inhibition is defined as the ability to stop (suddenly and completely) a planned or ongoing thought and action (Logan, 1994). This act of control is required in many real-life situations.

Some researchers, including Bunge and colleagues (2002) consider inhibition to be multifaceted construct. As such, inhibition has been distinguished between inhibition of an interference cue as labeled by Bunge and colleagues (2002) "interference suppression", and inhibition of habitual response as labeled by Bunge and colleagues (2002) "response inhibition". "Interference suppression" is when conflict is resolved by selectively attending to a relevant stimulus cue and ignoring a competing view for instance when we are at a cocktail party and we want to screen out all but one voice. "Response inhibition" is when conflict is resolved by refraining from executing an automatically-cued response (Bialystok & Viswanathan, 2009). Response inhibition is also present when the temptation to be resisted might be to indulge in pleasure when one should not (e.g. eating a rich dessert while trying to be on a diet). Another aspect of inhibition is when we need to be disciplined and stay on a task. In this current study I focus on the inhibition of habitual response or response inhibition as it is the type inhibition that plays a crucial role in language comprehension, as will be discussed later.

One of the challenges in understanding cognitive inhibition is the variety of tasks used to measure inhibition in children. Inhibition skill emerges around three years of age, but it develops significantly at the age of 7. If a child is not able to inhibit the irrelevant response he/she is considered to perseverate (i.e. repeat a response after the cessation of the original stimulus).

If we observe 3- to 4 year-olds we find that they may know perfectly well what they are supposed to do, and even try to do it, for example they understand that they should not touch the expensive vase on the shelf. But we may also find that some children still have difficulty inhibiting prepotent responses, i.e. suppressing their intention to touch the vase, due to an immature inhibition mechanism.

 By age 3 children demonstrate inhibitory control (Hughes, 1998) by showing ability to restrain their own actions. The ability to restrain their own actions provides that these children are capable of controlling their own behavior.

In the Snack delay task, which is a simple response inhibition task (see Garon et al., 2008), a child must delay the urge to eat a treat until the experimenter rings a bell. In this case, a child is required to withhold or delay an automatic response (that is, not

to eat the snack) until allowed. Whereas in complex response inhibition task, such as Day/night task, the child is supposed to hold a rule in mind, responding according to this rule and inhibiting a prepotent response. Task that measures inhibitory control in children older than 4 years old is the Simon says task in which a child must do what Simon asks him/her to do.

One of the most important paradigms of preschoolers is the delay of gratification paradigm (see Mischel, Ebbesen, Zeiss, 1973). In this paradigm there are two important types of measurements, the waiting and choice task. In the waiting task, also called simple response inhibition task (see Cragg & Nation, 2008), a child can delay eating a treat until being told to eat it. In this task Carlsson (2005) found that 50% of 24 months old could suppression eating a treat for 20 seconds, whereas3 years old suppressed the urge for 1 minute. This study also revealed that unlike these children, 4 years old were able to suppress eating their treat for 5 minutes. These findings show age improvements in children's inhibition skills.

The improvement of inhibition skills in children has been revealed in the choice version task of this paradigm, also called complex inhibition task (they involve larger amount of working memory). In this task a preschooler chooses between a small reward now and a larger reward then (Garon, 2008). Studies have found that choosing larger reward improved with age (i.e. a 5 year old showed better skills of suppression than the 3 year old) (e.g. Lemmon & Moore, 2001).

Age difference have been found in studies using the Stroop task which is a task where participants are required to name the color in which words are written rather than read the words. This task involves conflicting response since the ink color and word are discrepant and thus the child has to suppress or inhibit the dominant response to read the word. In a longitudinal study significant improvements in inhibiting conflicting responses were found between 24 months old and onward (e.g. Kochanska, Murrey & Koy, 1997).

Investigation of further improvement in inhibition beyond age of 5 has relied on the tasks that researchers used. As Best and Miller (2008) discuss, this type of game could have shown ceiling effect and as such produce finding of no inhibition improvement. However in one study conducted by Brocki and Bohlin (2004) using a Go-No-GO task, a significant improvement of inhibition between 7-9 years of age was found. In this task a child must respond only to "go" stimuli and exhibit responses to the "no-go" stimulus. Williams, Ponesse, Schachar, Logan and Tannock (1999) used a "stop-signal reaction time task" and found improvement of response inhibition up through age 12.

Response inhibition has been shown to be involved in reading abilities. Children with poor inhibitory mechanisms have difficulty in suppressing irrelevant information in the text. As a result, their memory capacity can become overloaded, which impairs their ability to maintain relevant information. Consequently, these children's integration, comprehension and memorization of text is much poorer than children with good inhibition skills (Borella, Caretti & Pelgrina, 2010). For a reader to understand the entire meaning of the text, he/she needs to maintain in working memory the most important information while inhibiting less important information. Thus, in some

cases poor comprehenders are readers who are poor at suppressing no-longer relevant information.

The next paragraphs explains the role of cognitive inhibition based on the CI model of word understanding. According to the CI model of word understanding, when compreheners first hear or read a word, information provided by that word activates various potential meanings. Then, constraints provided by lexical, semantic, syntactic, and other sources of information alter those meanings' levels of activation and eventually, one meaning becomes most strongly activated. That meaning is what comprehends access and incorporate into their mental structures (Kintsch, 1988). An individual with weak cognitive inhibitory skills may retain no longer-relevant information in working memory and/or allow irrelevant information to enter this workspace (Cain, 2009). Irrelevant information may interfere with the generation of the information to be processed and affect children's ability to do so. Moreover significant information could be poorly represented and disappear if the system is overloaded.

Poor comprehenders are less able to inhibit irrelevant information than their peers (De Beni & Paladino, 2000). According to the authors, improvement in reading comprehension that occurs in a critical phase of learning, from third to fourth grade when significant progress in reading comprehension is being made, appears to be explained by the efficiency of suppression mechanisms operating in working memory.

In another study Cain (2006) found that poor comprehenders did not differ from their peers with better comprehension on a short-term memory task that involved recalling short list of digits, concrete words, or abstract words. But when these individuals were given complex working memory tests that required them to supply final words in sentences and then remember those words later, poor comprehenders performed significantly more poorly than their peers with better comprehension. These findings were later supported by the study conducted by Schmidt (2010), examining 49 English native participants with two inhibitory tasks, a verbal task and a nonverbal task, which were used to assess participants' inhibitory control. There are no published studies that directly address the relation between poor comprehenders' higher order skills and their inhibition skills. As noted above, higher order skills such as inference making involve use of background information to make the meaning of the text coherent . If the background information compete between information that is no longer relevant to the comprehension of the text with the information that should be integrated in order to create the mental representation then inhibition plays a role.

Furthermore, in comprehension monitoring, which is considered to be an even higher order skill of comprehension the child is required to create adequate mental puzzles in order to create the mental representation during the online process of reading . Since this stage of comprehension involves monitoring one's own understanding and thinking we assume that weak inhibition skills would harm the child's ability to fully comprehend the story. Another research gap in this regard is examining whether inhibition skills contribute differently to lower level skills such as word decoding compared to higher level skills such as inference making and comprehension monitoring and control. Comparing three different age groups of children aged 5, 7 and 9 years would provide more data on how inhibition differs in terms of age as well .

4.2.   Updating

Updating is an executive function, which involves efficiently revising the components of working memory as new and relevant information becomes available (Carretti, Cornoldi, De Beni, & Romano, 2005). The updating function requires monitoring, coding and revising of incoming information for relevance by replacing old, no longer relevant information with newer, more relevant information (Morris & Jones, 1990).

The most common task used to measure updating function is the letter memory task (Morris & Jones, 1990). In the letter memory task (adapted from Morris & Jones, 1990), several letters from a list are presented per letter. The task is simply to recall the last 4 letters presented in the list. To ensure that the task requires continuous updating, the instructions required the participants to rehearse out loud the last 4 letters by mentally adding the most recent letter and dropping the 5th letter back and then saying the new string of 4 letters out loud (Miyake et al, 2000).

Many studies have considered updating to play a significant role in language comprehension (e.g. (Palladino, Cornoldi, DeBeni, & Pazzaglia, 2001)..

In the process of understanding a text a reader is engaged not only in maintaining information but also, for example, in merging that information with previous knowledge, therefore actively processing incoming information. Updating may be important for integrating new pieces of information from the text with previously read text to form coherent mental models. During the comprehension process, information that is no longer relevant needs to be suppressed or inhibited to prevent it from being integrated into the mental model and interfering with comprehension (Barnes et al., 2004). During the on-line process of reading the reader must reactivate information that was previously processed but no longer held in working memory by bringing it back into working memory from long term memory (Rapp & van den Broek, 2005). The updated information will help the readers to construct meaning of text. In particular the updated information should be essential for higher order skills such as making inferences, monitoring comprehension as well as controlling comprehension.

Furthermore research studies have found that updating is a specific aspect of working memory that mediates the relation between working memory and reading comprehension (Carretti, Cornoldi, De Beni, & Romano, 2005). Given the limits of working memory capacity, effective use of memory involves not only maintaining as much information as possible but also continuously selecting and updating this information. Failure in this process could lead to misinterpreting the content of a text (Blanc & Tapiero, 2001).

Palladino et al., (2001) investigated the process of updating with a task in which participants were asked to recall the three smallest objects from a list of words that was read to them. They administered a series of updating working memory tasks to groups of adolescents and young adults categorized according to their reading comprehension ability. To be successful, participants needed to retain an updated list by adding and dropping words according to the relative size of their real-life referents. Palladino et al. (2001) found that participants who were less skilled in reading comprehension recalled fewer objects in the correct order on the updating task than their better comprehending peers.

During the process of reading, the child should be able to reactivate information that was previously processed but no longer held in working memory. Thus this information needs to be brought back into working memory from-long term memory (Rapp & van den Broek, 2005). This reactivation of information into the working memory from long term memory is crucial for making inferences, as a result of which the mental model is constructed.

Miyake et al. (2000) argue that updating may involve "temporal tagging" to keep track of which information is old and no longer relevant by going beyond the maintenance and manipulation of information. The ability to update information that is held in mind develops later than simple retention. Updating develops later in infancy in the 2nd year, and there is continued development throughout the preschool period (Garon, Bryson & Smith, 2008).

Carreti et al (2005) conducted a study in which the updating task consisted of eight columns of pictures and eight corresponding lists of nouns. Columns of pictures consisted of 15 pictures, taken from a pool of 54 highly familiar and nameable pictures prepared for very young children. Each list of words included nine concrete nouns, all of which were displayed in the column of pictures.

To maintain the structure of the updating task used by Palladino and colleagues (2001), the list also included six smaller items. Group comparisons showed an impairment for poor coprehenders in carrying out the updating task. They recalled significantly fewer correct words and made significantly more intrusion errors than did good comprehenders. This result shows evidence in favor of the hypothesis that updating in working memory and updating in reading comprehension are related (Palladino, Carnoldi, DeBeni & Pazzaglia, 2001).This result suggests that memory updating does not involve only simple maintenance and substitution, but probably more general control and coordination of information activation/suppression in working memory. The results also offer an insight into the role of updating ability in reading comprehension. In fact, both the updating and the reading comprehension tasks require the participant to continuously update information and, when required, find appropriate information.

4.3. Cognitive flexibility

Cognitive flexibility is an executive function which involves being flexible enough to adjust to changed demands or priorities. It is the ability to update or change goal-oriented behavior in response to endogenous and exogenous changes in task goal. Flexibility in thinking is a vital competence in our everyday lives. It allows us to think of alternatives and to adapt to changes in our environment (Kloo et al, 2008). Cognitive flexibility is thought to be suberved by the other two, working memory and inhibition, and comes in much emerges later in development (Davidson et al. 2006). Some studies have shown the emergence of some simple flexible behaviors in infancy (e.g. Stahl & Pry, 2005). However it is well evidenced that the initial explicit switching ability in children occurs between the ages 3 and 5 years (Cragg & Chevalier, 2012). The switching ability emerges at around 7 years of age and it improves or in-

creases between the age of 7 and 9 years, in which period children show significant capacity in to dealing with multiple dimensions in switching tasks (Anderson, 2002).

Cognitive flexibility is assumed to emerge in children between the ages of 3 and 5 when using simple switching tasks (Espy, 1997). Children tend to make more errors in more demanding switching tasks. Research has shown that children from 3 years of age to 7 years of age have serious difficulties in switching between tasks which require maintenance of multiple dimensions (i.e. shape and then color etc).

Cognitive flexibility is the ability to changing perspectives, which requires inhibition of previous perspective and load into WM a different perspective (Diamond, 2013). Cognitive flexibility on children is measure by The Dimensional Change Card Sort (DCCS) (Frye et al., 1995; Zelazo, 2006). This is undoubtedly the most widespread task of such paradigms . In the first phase of the DCCS, children are explicitly instructed to sort cards according to a specific dimension (e.g., colour). In the second phase, children are required to switch and use the alternative dimension (shape) to sort the cards. It is now well documented that most 3-year-old children succeed in maintaining an initial sorting dimension over the first phase but fail at switching dimension in the second phase, instead perseverating on the initial criterion. However improvements are found in children 4- and 5-year-old who appear be able to switch easily from one task to another. Cognitive flexibility in adults is investigated using task-switching test the Wisconsin Card Sorting Task (Milner, 1964).This tasks helps the examiner measure the ease of switching between different sets of sorting rules. In this test each card can be sorted by color, shape, or number. The task for the participant is to deduce the correct sorting on the basis of feedback and to flexibly switch sorting rules whenever the experimenter gives feedback that the sorting criteria have changed.

Cognitive flexibility is involved in the acquisition of language that occurs during early childhood (e.g. Deak, 2003). In particular, studies have shown that pre-reading skills are related cognitive flexibility. Skilled readers must coordinate flexibly multiple aspects of reading tasks for successful comprehension (Cartwright, 2012).

An important aspect of learning to read is efficiency in accessing different kinds of linguistic information such as orthographic, phonological, and semantic information. The assumption that cognitive flexibility accounts for successful comprehension is supported by the idea formulated by Perkins and Unger (1999) in the following way: "Understanding a topic is a matter of being able to think and act creatively and competently with what one knows about the topic. In support of this argument, MacWhinney (1987) puts it, ''In order to learn a language, a child must have available a rich representational system and flexible ways of deciding between representations. (pp. 249–250).

Elementary school-aged children typically have difficulty attending flexibly to both semantic and phonological aspects of printed words (Cartwright, 2002). In an attempt to examine both graphophonological-semantic and general color-shape cognitive flexibility in beginning readers to explore the development of cognitive flexibility in reading skills Cartwright (2010) examined 68 first and second graders using color-shape flexibility task (Bigler & Liben, 1992; Cartwright, 2002), and graphophonological-semantic flexibility (Cartwright, 2002, 2007; Cartwright et al., 2006) . This study

revealed that cognitive flexibility is present at least as early as first grade and develop from first to second grades. Second graders performed significantly better than first graders on both tasks even when controlling for decoding skill.

Making inferences from text is a critical reading skill by which many new words are learned. Studies have shown that cognitive flexibility is crucial to the process of inference making in young children (e.g. Deak, 2003; Narasimham & Deak, 2001). This is mainly done by using the Flexible Induction of Meaning (FIM) test that requires children to infer the meanings of novel words by using changing predicate cues to flexibly shift attention among aspects of the referent (e.g. phrase : lives in a ….) to infer new word meanings. Although all children in this study had flexible thinking, it showed that the flexibility improved from 3 to 5 years of age.

Individual differences in flexible induction of (spoken) word meanings might predict abiding differences in children's ability to infer meaning from text (see Yuill & Oakhill, 1991). Testing cognitive flexibility in inference making during text comprehension in children, besides three other executive functions, would be an added value because it concerns controlled changes in cognitive activity over time, problems, or tasks.

Cognitive flexibility is considered to play a crucial role in languages with opaque orthography such as English (Cartwright, 2002). Opaque orthographies have a number of mappings between orthography and phonology which slows the development of word reading (Seymor et al., 2003). Thus, we would expect that this may require additional effort of children during the process of reading as they will have to make early use of the variety of information at their disposal (orthographic, phonological and semantic) and hence would account for additional flexibility on reading comprehension. French, Greek and Spanish all have has a more transparent orthographic system than English, and so do Greek and Spanish languages.

In the case of language with a more transparent orthographic system, there may be less need for them to return to other sources of information, raising the questions of whether flexibility us critical for early reading comprehension in these languages (Cole, Duncan & Blaye, 2014). In an attempt to answer these questions, Cole and colleagues (2014) examined sixty 7- year -old French children in a school in France. Besides comprehension tasks such as Pseudo-word decoding, word reading and passage reading comprehension, the flexibility task they used was those designed by Cartwright (2002), a) word flexibility and b) picture flexibility. Both tasks required simultaneous processing of two dimensions: phonology and semantics. Interestingly, the finding of this study showed that cognitive flexibility predicted reading skills in French language, although compared to English it has a more transparent orthography. However, considering that the authors tested only French speaking children, future studies including a comparison between French and other language with even higher orthographic transparency such as Greek, Spanish and or Albanian would add knowledge in this field.

Cole and colleagues (2014) extended this study to another experiment to test the hypothesis whether flexibility plays a role in word reading as a well as reading comprehension. They found that flexibility predicts second- grade reading from comprehension not only of texts but also of isolated words beyond the classic influence of

decoding skills. Future longitudinal studies examining the contribution of flexibility in word reading in different language would be an added value.

Besides the three core executive functions as conceptualized by Miyake et al (2000) in the current study the fourth executive function is included according to Baddeley and Hitch model (1974). The multicomponent model of working memory is considered to be important in this study as it provides vital information to not only suppressing and reactivated information but also to recalling previous information from the text.

4.4.   Executive Working Memory

The multicomponent model of working memory established by Baddeley and Hitch (1974) comprises three main mechanisms. The first is the central executive (or CE) which is involved in the control and regulation of the working memory system. The other two mechanisms include the phonologically based store (the phonological loop, or PL) and a visuospatial store, (the visuospatial sketchpad, or VS). The central executive's main role is to supervise the other two. It plays an important role in switching attention, and activating representations within long term memory. Unlike the two slave systems, the CE is not involved in temporary storage. The phonological loop and visuospatial sketchpad are specialized for the processing of information and temporary maintenance of information within both domains (verbally coded information and visually coded information). Short term memory is needed for temporary storage and integration of information, whereas long-term memory for more permanent storage and as a source of relevant knowledge (Oakhill & Garnham, 1988).

Working memory ability emerges in early infancy. The most common task used to assess holding in mind during infancy is the delayed response task. In this task, a toy is hidden at one of two possible locations. This task has shown that 6-month-old infants have capacity to hold in mind a representation over a delay. The capacity to retain a representation for a longer period of time appears by 12 months of age (Garon, Bryson & Smith, 2008).

For children above 2 years of age the task most widely used is the span task, which task assesses the phonological loop and visuo-spatial sketchpad. In this task a child is asked to repeat a list of digits or words. The span task can involve digits (digit span task), words (word span task) or objects (object span task) and the number of items retained increases with age (Gathercole, 1998). For all types of span tasks, working memory occurs after the preschool period.

Some other researchers argue that performance on complex WM (i.e. those tasks requiring a greater degree of processing such as the maintenance and manipulation of information) improves at least through adolescence (e.g. Best & Miller, 2010; Luna et al, 2004). Another task which is used to tap the ability to manipulate representations in mind is the backward span task. In this task a child is asked to repeat a list of digits or words backward, and it is mostly used with children above 3 years of age.

An issue of major concern in the process of measuring the working memory ability is the developmental relation between WM and inhibition. It seems that many complex inhibitory tasks place demands in working memory (Garon et al, 2008). As

such, tasks involving the combination of the two lead to greater difficulty for younger children (e.g. Carlson, 2005). The fact that the two functions are separable but still related to one another makes it difficult for the examiners to obtain a pure measure of one or another. Furthermore, according to some researchers (e.g. Engle 2002), working memory plays a critical role on cognitive flexibility (i.e. resistance to interference).

However, relatively little research has been done on more complex working memory tasks in young preschoolers.

The concept of working memory has been widely studied in relation to language comprehension, as according to Baddeley and Hitch (1974), WM enables humans to comprehend and retain information as well as acquire new knowledge. Readers with low working-memory capacity experience numerous constraints on how much information that can keep active as they read resulting in lower comprehension (Kendeou et al, 2014). Children with reading problems are found to have less short-term memory capacity than normal readers. In a study using a false recognition task, Kail et al, (1997) found that good comprehenders were able to manipulate and integrate information in working memory.

Working memory has been shown to serve as a source of relevant background knowledge which will be important in understanding text and making inferences, and problems will arise if children cannot use information from working memory to help them make sense of the text (Yuill & Oakhill, 1991). The importance of working memory in language comprehension is shown by a study of Sesma, Mahone, Levine, Eason and Cutting (2009) in which working memory was assessed with the Freedom from Distraction Scale of the Wechsler Intelligence Scale for Children. It was found that 9-15 year old poor comprehenders had significant deficits in working memory ["compared to good/average/better comprehenders"?].

In an attempt to find if poor comprehenders would show impaired working memory abilities, Pimperton and Nation (2010), examined 109 children aged 7-8 years using measure of reading skills and working memory. In this study working memory was assessed using two subtests from the Automated Working Memory Assessment (AWMA; Alloway, 2007), one measuring verbal working memory (listening recall) and the other measuring visospatial working memory (spatial recall). Data from this study confirmed previous hypothesis that poor comprehenders show impaired working memory. The ability to keep information is essential to higher level skills in language comprehension such as inference making. Working memory is at the core of the reader's ability to reflect on his or her understanding of the text. Thus, poor working memory results in inadequate inference making and comprehension monitoring (Kendeou et al, 2010).

Poor comprehenders may be less able to use semantic skills to support memory. Cain (2006) examined 9-10 years old by using a working memory sentence task and working memory counting task. It was found that the difference between poor comprehenders and good comprehenders was greater on the working memory counting task than on the sentence task. In addition, both groups recalled a greater number of items in the counting task than in the sentence task. Considering that this study involved only 26 children shows a small sample that includes the number of detecting

significant differences between poor and good comprehenders in their working memory capacity.

As one of the most important higher order skills, inference making has been examined in two groups of children 7-8 years old defined as skills comprehenders and less skilled comprehenders by Cain, Oakhill, Barnes and Bryant (2001). Working memory was used by literal questions that assess memory for information given literally in the text. The authors found that less skilled comprehenders may simply have poorer memory of the information necessary for inference generation. Less skilled comprehenders often failed to recall the information that had been integrated to generate inferences.

Using the Reading Span Test, Daneman and Carpenter (1980) observed that individual differences in working memory capacity are associated with individual differences in inference generation and text integration. The authors found that low-span readers were unlikely to identify the referent to a pronoun sufficiently when six or seven sentences intervened between two. In contrast, high-span readers could always identify the referent to a pronoun, regardless of the number of intervening sentences. These studies suggest that working memory capacity is a major determinant of the efficiency of various comprehension processes that are important in comprehension. Notwithstanding this fact, some researchers have shown that working memory is not the critical factor in comprehension, or at least not the only one (Perfetti, Landi & Oakhill, 2005). This argument was based on the finding from the study conducted by Cain and Oakhill (1999) in which study they showed that there were no differences at all between the literal memory (memory from the literal comprehension of the text) of children who are making inferences and those who are not.

In an attempt to examine the oral comprehension in children the contribution of WM in it, Chrysochoou & Bablekou (2011), through a cross-sectional study examined one hundred and eighty children aged from 5 to 9 years old in Greece. The aim of this study was to understand the relationship between WM (specifically, the PL and CE components) and oral comprehension in early childhood, based on the assumption that young children would have more to rely on the CE to execute demanding comprehension operations. The oral comprehension of children was assessed using a battery of five stories; the phonological loop was assessed with four tasks (a digit recall, a word list matching, a word listed recall task); and the CE was assessed using three tasks (listening recall, accounting recall, and backward digit span). The data showed that the CE made considerable contribution to oral comprehension in the preschool years, but the PL did not support higher order comprehension skills. In addition, comprehension control was the skill best predicted by the CE in five-year-olds and the verbal WM contribution to oral comprehension decreased with age. In this study it was indicated that future studies comparing the contribution of both verbal and visuospatial CE tasks could provide a clearer understanding of children's language comprehension skills. Furthermore, besides WM and its components contribution to language comprehension, the investigation of other executive functions would be a great asset in this field. In a similar study, Chrysochoou, Bablekou and Tsigilis (2011), in support of their hypothesis, found that the CE plays an important role when young children (8-10 years old) are faced with higher order comprehension skills. Interest-

ingly their finding showed that the reading fluency was not the mediator in the case of Greek language in contrast to evidence with English speaking children. Future studies investigating the relevant contribution of WM component in Albanian language as compared to English would be valuable in this field.

References

Alloway, T. P. (2007). Automated Working: Memory Assessment: Manual. Pearson.Anderson, R. C., & Pearson, P. D. (1988). A schema-theoretic view of basic processes in reading comprehension. Interactive approaches to second language reading, (pp. 37-55I.

Baddeley, A. (1996). The fractionation of working memory. Proceedings of the National Academy of Sciences, 93(24), 13468-13472.Baddeley, A. (2000). The episodic buffer: a new component of working memory?. Trends in cognitive sciences, 4(11), 417-423.

Baddeley, A. (2003). Working memory: looking back and looking forward. Nature Reviews Neuroscience, 4(10), 829-839.

Baddeley, A., & Hitch, G. (1974). Working memory. In G.A.Bower (Ed.),Recent advances in learning and motivation (Vol. 8, pp. 47–90). New York, NY: Academic Press..

Baddeley, A. (2007). Working memory, thought, and action. Oxford University Press.

Barnes, M. A., Huber, J., Johnston, A. M., & Dennis, M. (2007). A model of comprehension in spina bifida meningomyelocele: Meaning activation, integration, and revision. Journal of the International Neuropsychological Society, 13(05), 854-864.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. Journal of Educational Psychology, 74(4), 506.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. Child development, 81(6), 1641-1660.

Blanc, N., & Tapiero, I. (2001). Updating Spatial Situation Models: Effects of Prior Knowledge and Task Demands. Discourse Processes, 31(3), 241-262.

Borella, E., Carretti, B., & Pelegrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. Journal of Learning Disabilities, 43(6), 541-552.

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary reading skill differences in young adults. Journal of learning disabilities, 40(3), 226-243.

Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. Neuron, 33(2), 301-311.

Cain, K. (1996). Story knowledge and comprehension skill. Reading comprehension difficulties: Processes and intervention, 167-192.

Cain, K. (2009). Making sense of text: skills that support text comprehension and its development. Perspectives on Language and Literacy, 35(2), 11-14.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of educational psychology*, 96(4), 671.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. Journal of educational psychology, 96(1), 31, 31-42.

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. Reading and writing, 11(5-6), 489-503.

Cain, K., & Oakhill, J. (Eds.). (2008). Children's comprehension problems in oral and written language: A cognitive perspective. Guilford Press.

Carlson, S. M., Davis, A. C., & Leach, J. G. (2005). Less Is More Executive Function and Symbolic Representation in Preschool Children. Psychological Science, 16(8), 609-616.

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. Learning and Individual Differences, 19(2), 246-251.

Carretti, B., Cornoldi, C., De Beni, R., & Romanò, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. Journal of experimental child psychology, 91(1), 45-66.

Carroll, J. B. (1993). Human cognitive abilities: A survey of factor-analytic studies. Cambridge University Press.

Cartwright, K. B. (2002). Cognitive development and reading: The relation of reading-specific multiple classification skill to reading comprehension in elementary school children. Journal of Educational Psychology, 94(1), 56.

Cartwright, K. B. (2012). Insights from cognitive neuroscience: The importance of executive function for early reading development and education. Early Education & Development, 23(1), 24-36.

Chevalier, N., & Blaye, A. (2009). Setting goals to switch between tasks: Effect of cue transparency on children's cognitive flexibility. Developmental Psychology, 45(3), 782.

Chrysochoou, E., & Bablekou, Z. (2010). Phonological loop and central executive contributions to oral comprehension skills of 5.5 to 9.5 years old children. *Applied Cognitive Psychology*, 25(4), 576-583.

Chrysochoou, E., Bablekou, Z., & Tsigilis, N. (2011). Working memory contributions to reading comprehension components in middle childhood children. *The American journal of psychology*, 124(3), 275-289.

Colé, P., Duncan, L. G., & Blaye, A. (2014). Cognitive flexibility predicts early reading skills. Cognitive Science, 5, 565.

Cragg, L., & Chevalier, N. (2012). The processes underlying flexibility in childhood. *The Quarterly Journal of Experimental Psychology*, 65(2), 209-232.

Cutting, L. E., Materek, A., Cole, C. A., Levine, T. M., & Mahone, E. M. (2008). Effects of fluency, oral language, and executive function on reading comprehension performance. Annals of dyslexia, 59(1), 34-54.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, 19(4), 450-466.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. Neuropsychologia, 44(11), 2037-2078.

Deak, G. O. (2003). The development of cognitive flexibility and language abilities. Advances in child development and behavior, 31, 273-328.

Deák, G. O., & Narasimham, G. (2003). Is perseveration caused by inhibition failure? Evidence from preschool children's inferences about word meanings. *Journal of Experimental Child Psychology*, 86(3), 194-222.

De Beni, R., & Palladino, P. (2000). Intrusion errors in working memory tasks: Are they related to reading comprehension ability?. Learning and Individual Differences, 12(2), 131-143.

Diamond, A. (2013). Executive functions. Annual review of psychology, 64, 135-168.

Dixon Jr, W. E., & Smith, P. H. (2000). Links between early temperament and language acquisition. Merrill-Palmer Quarterly (1982-), 417-440.

Dunn, D. M., & Dunn, L. M. (1959). Peabody Picture Vocabulary Test—Third Edition. 1959-2007; published by Pearson, 19500 Bulverde Road, San Antonio, TX 78259.

Dunn, L. M., & Markwardt, F. C. (1970). Peabody individual achievement test. American Guidance Service.

Eason, S. H., & Cutting, L. E. (2009). Examining sources of poor comprehension in older poor readers. Beyond decoding: The behavioral and biological foundations of reading comprehension, 263-286.

Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader–text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of educational psychology*, 104(3), 515.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3), 339-350.

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. Developmental neuropsychology, 26(1), 465-486.

Farrar, M. J., & Ashwell, S. Y. L. V. I. A. (2008). The role of representational ability in the development of phonological awareness in preschool children. Literacy processes: Cognitive flexibility in learning and teaching, 71-89.

Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and language*, 27(3), 235-244.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. Cognitive Development, 10(4), 483-527.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. Psychological bulletin, 134(1), 31

Gardner, H. (1985). Frames of mind: The theory of multiple intelligences. Basic books.

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry*, 39(1), 3-27.

George, Marie St, Suzanne Mannes, and James E. Hoffman. "Individual differences in inference generation: An ERP analysis." *Journal of Cognitive Neuroscience* 9.6 (1997): 776-787.

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: a component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(2), 245.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3), 430.

Goldman, S. R., Varma, S., & Cote, N. (1996). Extending capacity-constrained construction integration: Toward "smarter" and flexible models of text comprehension. Models of understanding text, 73-113.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. Psychological review, 101(3), 371.

Goswami, U. (2001). Early phonological development and the acquisition of literacy. Handbook of early literacy research, 1, 111-125.

Hacker, D. J. (1997). Comprehension monitoring of written discourse across early-to-middle adolescence. Reading and Writing, 9(3), 207-240.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, 16(2), 233-253.

Jenkins, J. R., Fuchs, L. S., Van Den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, 95(4), 719.

Johnston, A. M., Barnes, M. A., & Desrochers, A. (2008). Reading comprehension: Deveopmental processes, individual differences, and interventions. Canadian Psychology/Psychologie canadienne, 49(2), 125.

Kendeou, P., Broek, P., Helder, A., & Karlsson, J. (2014). A Cognitive View of Reading Comprehension: Implications for Reading Difficulties. Learning Disabilities Research & Practice, 29(1), 10-16.

Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. (2007). Comprehension in preschool and early elementary children: Skill development and strategy interventions. Reading comprehension strategies: Theories, interventions, and technologies, 27-45.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. Psychological review, 95(2), 163.

Kintsch, W. (1998). Comprehension: A paradigm for cognition. Cambridge university press.

Kloo, D., Perner, J., Kerschhuber, A., Dabernig, S., & Aichhorn, M. (2008). Sorting between dimensions: Conditions of cognitive flexibility in preschoolers. *Journal of Experimenta l Child Psychology*, 100(2), 115-134.

Lorsbach, T. C., Wilson, S., & Reimer, J. F. (1996). Memory for relevant and irrelevant information: Evidence for deficient inhibitory processes in language/learning disabled children. Contemporary educational psychology, 21(4), 447-466.

Luna, B., & Sweeney, J. A. (2004). The emergence of collaborative brain function: FMRI studies of the development of response inhibition. Annals of the New York Academy of Sciences, 1021(1), 296-309.

Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. Child development, 643-655.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, 91(1), 3.

Miyake, A., Friedman, N., Emerson, M. J., & Witzki, A. H. (2000). The unity and diversity of executive functions and their contributions to complex ''frontal lobe'' tasks: A latent variable analysis. Cognitive Psychology, 41, 49–100.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British journal of psychology*, 81(2), 111-121.

Munakata, Y. (2001). Graded representations in behavioral dissociations. Trends in cognitive sciences, 5(7), 309-315.

Oakhill, J., & Cain, K. (2007). Introduction to comprehension development. Children's comprehension problems in oral and written language: A cognitive perspective, 3-40.

Oakhill, J., Cain, K., & Yuill, N. (1998). Individual differences in children's comprehension skill: Toward an integrated model. Reading and spelling: Development and disorders, 343-367.

Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. Reading and Writing, 23(2), 189-208.

Oakhill, J., & Garnham, A. (1988). Becoming a skilled reader. Basil Blackwell.

Olson, R., Forsberg, H., Wise, B., & Rack, J. (1994). Measurement of word recognition, orthographic, and phonological skills.

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. Memory & cognition, 29(2), 344-354.

Pany, D., Jenkins, J. R., & Schreck, J. (1982). Vocabulary instruction: Effects on word knowledge and reading comprehension. Learning Disability Quarterly, 5(3), 202-215.

Paris, S. G., & Myers, M. (1981). Comprehension monitoring, memory, and study strategies of good and poor readers. *Journal of Literacy Research*, 13(1), 5-22.

Perkins, D. N., & Unger, C. (1999). Teaching and learning for understanding. Instructional-design theories and models: A new paradigm of instructional theory, 2, 91-114.

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. The neurocognition of language, 167-208.

Perfetti, C. A. (1985). Reading ability. Oxford University Press.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The Acquisition of Reading Comprehesion Skill.

Pimperton, H., & Nation, K. (2010). Suppressing irrelevant information from working memory: Evidence for domain-specific deficits in poor comprehenders. *Journal of Memory and Language*, 62(4), 380-391.

Popper, K. (1978). Natural selection and the emergence of mind. Dialectica, 32(3-4), 339-355.

Rapp, D. N., Broek, P. V. D., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. Scientific Studies of Reading, 11(4), 289-312.

Roch, M., Florit, E., & Levorato, C. (2011). Follow-up study on reading comprehension in Down's syndrome: the role of reading skills and listening comprehension. *International Journal of Language & Communication Disorders*, 46(2), 231-242.

Sénéchal, M., Thomas, E., & Monker, J. A. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology*, 87(2), 218.

Sesma, H. W., Mahone, E. M., Levine, T., Eason, S. H., & Cutting, L. E. (2009). The contribution of executive skills to reading comprehension. Child Neuropsychology, 15(3), 232-246.

Seigneuric, A., Ehrlich, M. F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. Reading and writing, 13(1-2), 81-103.

Shankweiler, D. (1999). Words to meanings. Scientific Studies of Reading, 3(2), 112-127.

Shefelbine, J. L. (1990). Student factors related to variability in learning word meanings from context. *Journal of Literacy Research*, 22(1), 71-97.

Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. Perspectives on individual characteristics and foreign language education, 6, 27.

Snow, C. (2002). Reading for understanding: Toward an R&D program in reading comprehension. Rand Corporation. Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: evidence from a longitudinal structural model. Developmental psychology, 38(6), 934.

Snowling, M. J. (2005). Literacy outcomes for children with oral language impairments: Developmental interactions between language skills and learning to read. The connections between language and reading disabilities, 55-75.

Stahl, L., & Pry, R. (2005). Attentional flexibility and perseveration: Developmental aspects in young children. Child Neuropsychology, 11(2), 175-189.

Stanovich, K. E. (2000). Progress in understanding reading: Scientific foundations and new frontiers. Guilford Press.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolersA follow-up into adolescence. *Journal of Speech, Language, and Hearing Research*, 41(2), 407-418.

Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. Learning Disabilities Research & Practice, 15(1), 55-64.

van den Broek, P. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. Comprehension processes in reading, 423-445.

van den Broek, P., Risden, K., Fletcher, C. R., & Thurlow, R. (1996). A "landscape" view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. Models of understanding text, 165-187.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). Comprehensive test of phonological processing: CTOPP. Austin, TX: Pro-ed.

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. Reading and Writing, 16(1-2), 5-20.

Wiederholt, J. L., & Bryant, B. (2001). Gray oral reading test-(GORT-4). Austin, TX: Pro-ed.

Woodcock, R. W. (1987). Woodcock reading mastery tests, revised. Circle Pines, MN: American Guidance Service.

Yuill, N., & Oakhill, J. (1991). Children's problems in text comprehension: An experimental investigation. Cambridge University Press.

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): A method of assessing executive function in children. Nature Protocols, 1(1), 297-301.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. Psychological bulletin, 123(2), 162.

Local Music Scenes. Forms of musical creation, production and consumption in Thessaloniki, Greece

Nikos Dallas

PhD Candidate in Sociology of Arts
School of Journalism & Mass Media Communication,
Aristotle University of Thessaloniki, Greece
dallasland@hotmail.com

Abstract. Collective music practices have great relevance for intimate and social realm within a certain locale affecting a vast number of people, either as participants and producers, or as audiences, and the way they experience urban life as a whole. Adopting the concept of 'scene' and treating its related theories with a critical view, I examine music activities around a particular music genre, 'experimental music', in Thessaloniki, Greece. Through participant observation and in-depth, semi-structured interviews, I explore how individuals make use of cultural resources in the present socio-economic conjuncture and what are the consequences for the city itself.

Keywords: *music scenes, locality, collective practices, urban life*

1   Introduction

In his stimulating book *Why Music Matters* (2013), David Hesmondhalgh examines the social value of music by exploring the dual capacity of music to allow the private self to flourish, and to offer collective experiences to thousands and sometimes millions of people. In an era that celebrates individuality and personal autonomy to such a great extent – even music technology itself does so by providing for a long time products like walkman, MP3 player and iPod that allow listeners to cut themselves off from their surroundings – the public dimension of musical experience, from attending live performances at gigs and festivals to participating in groups of shared musical tastes of any kind and size, is one major factor in socializing and constructing collective identities in modern societies. Yet, another equally important factor that has particularly emerged in the 1990s, that is 'the years of globalization' and the proliferation of internet, and has been strongly reinforced during the present socio-economic crisis, is locality, the reference to a particular geographic and sometimes demographic area ranging in size from a small neighbourhood to a whole metropolitan city. These two elements – collective musical practices and the locale – and the way they interact with each other are central in my research that focuses on how music is linked to sociality and community and how participation in musical

activities affects both people's lives in a certain place and the formation of this place per se. On the one hand, I examine the structure of local musical groupings by emphasizing on the personal characteristics of their members (age, gender, educational and economic background), the reasons, needs and motives that lead them to such activities, and the external aspects and circumstances that form and define the latter (socio-economic milieu, state and local policies, cultural industries); on the other hand, I explore the way that the locale impacts on these groupings in reference to its historical continuity and tradition, its cultural and social structures and its public and private institutions. Music can become an extremely useful key in this process, since it is considered to be the most prominent marker of differentiation among all other forms of cultural expression (Bourdieu, 1979) and perhaps the cultural product that crosses boundaries and frontiers most frequently (Frith, 1996).

2   Theorizing local music scenes

Although music played no significant part in the original Centre for Contemporary Cultural Studies (CCCS) analysis (Laing, 1985), the tradition in subcultural studies is a primary theoretical tool for approaching collective musical practices, even if it is just to criticize the body of work that has been produced by the scholars of this particular Department of the University of Birmingham in the second half of the 1970s. CCCS theorists have adopted from the Chicago School the concept of 'subculture' (Bennett, 2006: 222) in order to study how the post-war British working-class youth groups such as the teddy boys, mods, skinheads and punks reproduced, negotiated and transformed their material conditions through signifying cultural practices. Grounded in Marxist and neo-Marxist theories of class conflict and strongly influenced by Gramsci's 'hegemony', Althusser's 'relative autonomy' and Barthes' and Levi-Strauss' 'bricolage', thinkers of the so called 'Birmingham School' suggested that the different subcultures were addressing the 'class problematic' of the particular strata from which they were drawn and at the same time they were providing highly stylised and symbolic, and therefore imaginary, solutions to their members (Clarke *et* al., 1975: 35).

   The impact of the CCCS approach has been tremendous and although it was never a unified set of ideas, it became the object of stern and often levelling critique, focused on its tendency to view (white, male, heterosexual) working-class youth subcultures as potentially resistant in political terms (Hodkinson & Deicke, 2007) and its members as a stable and coherent group of social agents who acted in a deterministic way. This kind of criticism was also combined with a confrontation for putting class rather than gender, race or/and sexuality at the centre in theorizing the lives of young people and with a plea for more grounded, empirical research (Muggleton, 1997: 167). 'Decentring the Centre', to use Maureen McNeil's play on words (Turner, 2003: 62), became a duty and in some cases an obsession for subsequent theorists in cultural studies who tended to overlook – sometimes deliberately – that the youth subcultures project was formed at a particular historical, cultural and political conjuncture, as referenced in the subtitle of Clarke and colleagues' classic work *Resistance through Rituals: Youth Subcultures in Post-war Britain* (Griffin, 2011).

But as subcultural approach was a product of its time, so was the body of criticism that emerged during the following years. Writing in the wake of postmodernism 'post-subcultural' scholars, with great respect to the work of Weber and Baudrillard, challenged the CCCS class-based subcultural approach, arguing that it could not capture the fluidity of the new neo-liberal, individualistic environment, and attempted to re-label their contemporary less structurally bounded groupings. Besides, according to Steve Redhead (1997), a key advocate of 'post-subculturalism', sometime between the advent of punk in the 1970s and the emergence of rave culture in the 1980s, 'the moment of "subculture"' passed into history (Hall & Jefferson, 2006: xix). In this context, various concepts emerged as offering new ways of conceiving cultural collectivities from 'tribes' or 'neo-tribes' (Maffesoli, 1996) to 'lifestyles' (Bennett, 1999) and 'milieu cultures' (Webb, 2007), and from 'taste cultures' (Lewis, 1992) and 'articulation' (Hesmondhalgh, 2005) to 'scenes' (Straw, 1991).

Before being introduced in the academic discourse by Will Straw in the early 1990s, the term 'scene' – almost exclusively referred to musical activities – was originally used primarily in everyday and journalistic discourses especially in association with the demiworld of jazz in the 1940s (Bennett & Richardson, 2004). In his influential article *Systems of Articulation, Logics of Change: Communities and Scenes in Popular Music*, Sraw defines a musical scene as 'that cultural space in which a range of musical practices coexist, interacting with each other within a variety of processes of differentiation, and according to widely varying trajectories of change and cross-fertilisation' (1991: 372). In this way Straw disengages collective musical activities from the more fixed unities of subcultures and distinguishes scene's ability to evoke both the cozy intimacy of community and the fluid cosmopolitanism of urban life (Straw, 2001: 248).

Thereupon the notion of scene has increasingly been used as a model for academic research on the creation, production and consumption of popular music. Two notable examples are Shank's (1994) study of the rock scene in Austin, Texas and Bennett and Peterson's (2004) work on music scenes in general. The former suggests that a 'scene' is 'an overproductive signifying community, arising from a certain intensity of commitment' which (this commitment, that is) is the 'necessary condition' for the production of exciting music within a place (Shank, 1994: 122). The latter develop a three-tier model of 'scenes' consisting of local, trans-local and visual scenes arguing that no one music scene need be exclusively local, trans-local or visual but may exhibit properties from all three of these categories.

I am using the concept of 'scene' because it both avoids misleading overtones of concreteness and stability, and allows for a more varied range of collective musical participation. Actually, even when the term is used to connote a cluster of musical activities taking place in a geographically bounded space ('local scene') neither is limited to face-to-face gatherings and contacts nor does premise a spectacular visual attire of its members. Furthermore, 'scene' serves as a very useful analytical tool for studying musical practices not only because it is strongly associated with music but also because of its extensive use and appropriation in everyday contexts which help the researcher contact the agents of this particular field. In any case, the concept of 'scene' and its related theories are treated with a critical view and an intention to

understand people as individuals able to make use of cultural resources within a certain locale, and also to understand all those factors – social, economic and political – that affect and sometimes determine and limit people's choices, tastes and actions.

3   Framework and methodology

My research takes place in Thessaloniki, the second largest city in Greece whose history spans some 2,300 years. Situated in the centre of the Balkans, the city used to house different ethnicities and nationalities that created a multicultural and cosmopolitan character which started to fade out from the second half of the 20th century onwards. Today little remains from this earlier cosmopolitan era which was followed by a period of conservatism and introversion expressed for many years by the local political and religious authorities and elites, and nowadays marked by the recession of the current socio-economic crisis. However, Thessaloniki has always been a city of cultural significance, appraised for its contribution to arts and literature and also included in various lists of 'the best cities' for nightlife and entertainment. The latter has been a sector that was overemphasized during the 1990s and persistently supported and raised by the media.

In the field of music, Thessaloniki has a long and rich tradition and legacy reflected in numerous public and private institutions associated with musical activities, including the School of Music Studies of the Aristotle University, conservatoires, choirs, symphony orchestras, recording and rehearsal studios, record companies, venues for gigs and live performances and so on. As far as popular music is concerned the domination of 'laïkó' (urban-folk) and 'pop', a Westernised genre with Greek lyrics (Tsioulakis, 2011: 182), is unquestionable but there is always space for other genres to flourish, thanks to micro-communities of musicians, producers and audiences – 'scenes' – that are ranged in homogeneity, size, duration and public visibility. For instance, back in the mid-1980s a vital rock scene emerged in Thessaloniki with a lot of fans and bands, some of which had great creative and commercial success, such as Trypes and Morá Sti Fotiá, while there have been other scenes, much smaller in size and far less famous in terms of the recording industry, like the one of Goths which has its own bearers, bands (IAMBIA) and venues (X-club, Eightball), and its own kind of success.

Living in Thessaloniki since 1993 and having a keen interest in music, I have been a witness to the emergence, alteration and decline of a number of scenes which influenced in diverse and various ways local music-making and urban-living as well. I would not call myself a member of any particular scene, not at least in terms of subcultural commitment, but on the other hand, as Kahn-Harris (2000: 25) suggests writing about the Extreme Metal scene, 'even the simple action of buying a CD means to become "involved" in a scene, in however slight a way, by virtue of causing some sort of effect within it'. In this light, I have been participating as part of the audience in collective activities and experiences related to experimental music for several years.

Although the term 'experimental music' has a very specific meaning in the world of music academia referring to a compositional tradition which arose in the mid-20th

century, applied particularly in North America to music composed in such a way that its outcome is unforeseeable, and whose most famous and influential exponent was John Cage (Grant, 2003: 174), I use this term in a much broader sense; I am referring to music within specific genres – mainly, but not exclusively, in electronic music – that pushes against their boundaries and definitions by using new forms of instrumentation (such as electronics or novel instruments) or traditional instruments played in new ways, and by incorporating unorthodox, new, distinctly unique ingredients. Of course, this is not a bounded approach or an attempt for a once-and-final definition (one could suggest, for example, that there is some degree of experimentation almost in any musical composition) but the term is widely used and appropriated in this sense both by musicians and fans, and the recording industry and broader social networks. Besides, as Keith Negus (2002) warns new bands, 'any attempt to refuse musical labeling will not help you get bookings'.

Three other parameters that are pivotal in my research are (i) that it concerns both amateur and professional musicians, (ii) that it does not focus exclusively on youth and (iii) that it examines the effect not only of the recording industry in local music-making but of the DIY (do-it-yourself) ethics as well. As Finnegan (1989: 15) rightly argues, one of the interesting characteristics of local music organization is the absence of an absolute distinction between 'the amateur' and 'the professional', as neither occupational nor educational criteria can provide a clear dividing line between the two. On the other hand, although scholars for many years used to take for granted the equation of legitimate popular music fandom with youth, contemporary music scenes become so increasingly multi-generational (Bennett & Taylor, 2012) that call us to reconsider the concept of ageing. To put it in Bennett's words 'where investment in a musical style has been particularly intensive during one's teenage or 20-something years such investment may continue past 30, into middle age and perhaps later life' (2006: 221). Finally, although the recording industry – besides its transformations – still determines musical products and their meanings, collapse of the Greek economy reinforces DIY production which has a long tradition in collective musical practices and, with the support of new technological advances and the growing interest in live music (Frith, 2007), encourages everyone to take some action, rejecting the bifurcation of production into material and immaterial dimensions (Luvaas, 2013: 130).

Attempting to map the relationship between local music-making processes and the broader social processes which inform both cultural and everyday experiences in an urban setting, my methodology is based first of all on participant observation, involving the systematic description of events, behaviours and artefacts in the situation under study (Marshall & Rossman, 1989: 79) and developing a holistic understanding of the particular field (DeWalt & DeWalt, 2002: 92). In conjunction with this qualitative method of data collection that has been widely used in the study of music, I am going to conduct a series of in-depth, semi-structured interviews with exponents of all the agents involved in experimental musical practices, that is musicians, producers and fans, in order to gain an insight into individual opinions and evaluations. Holding a degree of insider status, is a good condition, as Hodkinson observes, for the achievement of successful and productive interactions with

participants (2005: 136), so I am trying to take advantage of it and not just to display an uncritical acceptance of insider knowledge as an end in itself, as a number of recent empirically focused sociological studies do (Bennett, 2002: 456). As Finnegan says, 'being *too* much of an insider (and ceasing to be a detached observer) [is] always a danger' (1989: 343).

4   Conclusions

In spite of the profound problems we face in our everyday lives because of the current socio-economic crisis that has particularly negative impacts on urban areas, musical activities continue to provide a basis both for self expression and intimate relations with others, and for collective, public practices and experiences. Participating in flexible groupings of musicians, producers and audiences that realise their common musical tastes within a certain place via mediated and face-to-face communication might contribute not only to socializing and entertaining – which are certainly very important aspects of private and social life – but also to understanding our ability to join together with other people in experiences of community and sociability. But apart from shaping our sense of who we are and how we might live together in contemporary societies, it might also contribute to reinventing the locale by decoding the tropes and typologies of post-globalisational processes in the social and cultural realm. In this light, local music scenes might be a site where participants can redefine themselves and the identity of their city, and simultaneously a factor that can affect both cultural policymaking and regulation at local, regional and even national level, and the operation of cultural industries and institutions. Besides, the relationship between music-making and the locale continues to be an extremely thinly mapped field in Greek research that needs to be explored so that it can no longer be seen as secondary to the 'real' social relations; as Bennett and Taylor accurately observe (2012: 231) in this late modern social context, identities are argued to be the product of the everyday appropriation and re-inscription of mass-produced cultural resources with new meanings drawn from localised experiences.

References

Bennett, A. (1999) Subcultures or neo-tribes? Rethinking the relationship between youth, style and musical taste. *Sociology*. Vol 33, no 3, 599-617

Bennett, A. (2002) Researching youth culture and popular music: a methodological critique. *British Journal of Sociology*. Vol 53, no 3, 451-466

Bennett, A. & Peterson, R. (eds) (2004) *Music Scenes: Local, Translocal and Virtual*, Vanderbilt University Press, Nashville

Bennett, A. (2006) Punk's Not Dead: The continuing significance of punk rock for an older generation of fans. *Sociology*. Vol 40, no 2, 219-235

Bennett, A. & Taylor, J. (2012) Popular music and the aesthetics of ageing. *Popular Music*. Vol 31, no 2, 231-243

Bourdieu, P. ([1979] 2006) *Distinction: A Social Critique of the Judgment of Taste*, Harvard University Press, Cambridge

Clarke, J. – Hall, St. – Jefferson, T. & Roberts, Br. ([1975] 2006) Subcultures, cultures and class, in Hall, St. & Jefferson, T. (eds), *Resistance Through Rituals. Youth subcultures in post-war Britain*, Routledge, Oxon, 3-59.

DeWalt, K. & DeWalt, B. (2002) *Participant observation: a guide for fieldworker*s, AltaMira Press, Walnut Creek, CA

Finnegan, R. ([1989] 2007) *The Hidden Musicians: Music Making in an English Town*. Middletown, Wesleyan University Press, Connecticut

Frith, S. (2007) Live Music Matters. *Scottish Music Review*. Vol 1, no 1, 1-17

Grant, M. J. (2003) Experimental Music Semiotics. *International Review of the Aesthetics and Sociology of Music*. Vol 34, no 2, 173-191

Griffin, Chr. E. (2011) The trouble with class: researching youth, class and culture beyond the 'Birmingham School'. *Journal of Youth Studies*. Vol 14, no 3, 245-259

Hall, St. & Jefferson, T. (2006) Once more around *Resistance through Rituals*, in Hall, St. & Jefferson, T. (eds), *Resistance Through Rituals. Youth subcultures in post-war Britain*, Routledge, Oxon, vii-xxxii.

Hesmondhalgh, D. (2005) Subcultures, Scenes or Tribes? None of the above. *Journal of Youth Studies*. Vol 8, no 1, 21-40

Hesmondhalgh, D. (2013) *Why Music Matters*, Wiley-Blackwell, West Sussex

Hodkinson, P. (2005) 'Insider Research' in the Study of Youth Culture. *Journal of Youth Studies*. Vol 8, no 2, 131-149

Hodkinson, P. & Deicke, W. (eds) (2007) *Youth Cultures: Scenes, Subcultures and Tribes*, Routledge, New York-London

Kahn-Harris, K. (2000) 'Roots'?: the relationship between the global and the local within the Extreme Metal scene. *Popular Music*. Vol 19, no 1, 13-30

Laing, D. (1985) *One Chord Wonders. Power and meaning in Punk Rock*, Open University Press, Milton Keynes

Lewis, G. H. (1992) Who do you love?: The dimensions of musical taste. In *Popular Music and Communication* (J. Lull ed.), Sage, London

Luvaas, Br. (2013) MATERIAL INTERVENTIONS: Indonesian DIY Fashion and the Regime of the Global Brand. *Cultural Anthropology*. Vol 28, no 1, 127-143

Maffesoli, M. (1996) *The Time of the Tribes: The Decline of Individualism in Mass Society*, Sage, London

Marshall, C. & Rossman, Gr. B. (1989) *Designing qualitative research*, Sage, Newbury Park CA

Muggleton, D. (1997) The Post-subculturalist, in Redhead, St., Wynne, D. & O'Connor, J. (eds), *The Clubcultures Reader: Readings in Popular Cultural* Studies, Blackwell, Oxford, 185-203.

Negus, K. (2002) *Music Genres and Corporate Cultures*, Routledge, London-New York

Redhead, St. (1997) *Subculture to Clubcultures: An Introduction to Popular Cultural Studies*, Blackwell, Oxford

Shank, B. (1994) *Dissonant Identities: The Rock 'n 'Roll Scene in Austin, Texas*, Wesleyan University Press, London

Straw, W. (1991) Systems of Articulation, Logics of Change: Communities and Scenes in Popular Music. *Cultural Studies*. Vol 53, 368-388

Straw, W. (2001) Scenes and Sensibilities. *Public*. Vol 22-23, 245-257

Tsioulakis, I. (2011) Jazz in Athens: Frustrated Cosmopolitans in a Music Subculture. *Ethnomusicology Forum*. Vol 20, no 2, 175-199

Turner, Gr. (2003) *British Cultural Studies. An introduction*, Routledge, London

Webb, P. (2007) *Exploring the Networked Worlds of Popular Music: Milieu Cultures*, Routledge, New York-London

Adaptation of tropical urban park management in Kuala Lumpur, Malaysia towards delivering ecologically sustainable landscape practice: Comparing the benefits and challenges

Roziya Ibrahim[1], Andy Clayden [2], and James Hitchmough[3]

[1] Department of Landscape, the University of Sheffield, UK
`arp11ri@sheffield.ac.uk`

Abstract. Since the early 19th century, urban parks have evolved to fulfil the social needs of urban communities. Today, these managed urban landscapes have a more important role in providing ecosystem services to cities, which are currently facing environmental degradation, in response to urbanization and climate change. Surprisingly, urban parks in Malaysia are not being used to overcome these problems. It becomes a considerable burden to local authorities in adapting to these challenges, as current maintenance is still based on traditional horticultural practice, which consumes high resources. However, there is a possibility that tropical urban parks can be managed more ecologically sustainably. This paper aims to highlight the potential for tropical urban park management in Kuala Lumpur, Malaysia to change their approach towards one of more ecologically sustainable practice by learning from the experience of developed countries. However, considering the climate and cultural differences, a comparison of ecologically sustainable practices between temperate and tropical contexts is necessary, in search of the practices most adaptable to Kuala Lumpur. This paper draws on data from interviews with the respective landscape architects and park managers and site observations from the selected precedent studies in the UK and Singapore. It outlines and discusses the potential benefits and challenges of changing the management practice into ecologically sustainable practice. The paper serves as guidelines for design and management alternatives that could inform and foster ecologically sustainable practice for tropical urban parks in Kuala Lumpur, considering the local context and background.

Keywords: *Urban Park Management, Ecologically Sustainable practice, Tropical Urban Park*

1   Introduction

Globally, the changing roles and functions of urban parks have a strong influence on their design and management, which leads to remarkable changes in their landscape (Zipperer & Zipperer, 1992). Simultaneously, this changing landscape has significantly influenced the way people perceive the new environments created (Antrops, 2005). Furthermore, Cranz & Boland (2004) argue that in fulfilling the demands of the 21st century, changes in urban parks should not only respond to social needs, but also fulfil the ecological needs of the urban environment, without alienating the park users (Gobster et al., 2007), especially through their aesthetic appearance (Cranz & Boland,

2004; Gobster et al.,2007). Currently, there is a strong movement toward a more ecologically sustainable landscape practice in developed countries. However, despite their benefits in providing ecosystem services for the city (Lovell & Johnston, 2008), it is also recognized that this approach may create challenges for those involved in managing these landscapes (Calkins, 2005) and also the public (Nassauer, 1995). It is important to understand how the tropical urban parks in Kuala Lumpur, Malaysia inherited a similar process of evolution and translated it to their own context, in order to identify potential for changing towards more ecologically sustainable practice.

2   The colonial influence on the evolution of urban parks in Kuala Lumpur, Malaysia

Historically, urban parks have evolved to adapt to social changes and demands. According to Jordan (1994), the industrial revolution in the early 19th century sparked the public park movement in the United Kingdom, resulting in more public parks being developed in urban areas. For example, the 125 acre Birkenhead Park, on the Wirral Peninsula, built in 1847, was the first publicly-funded municipal park in the United Kingdom (Conway, 2000). Later, the design of this park inspired Frederick Law Olmstead to adopt a similar approach in the larger scale development of Central Park, New York in 1857 (Lawrence, 2008). Both parks have become precedents for development of urban parks in their respective countries as well as other parts of the world, including Kuala Lumpur, Malaysia.
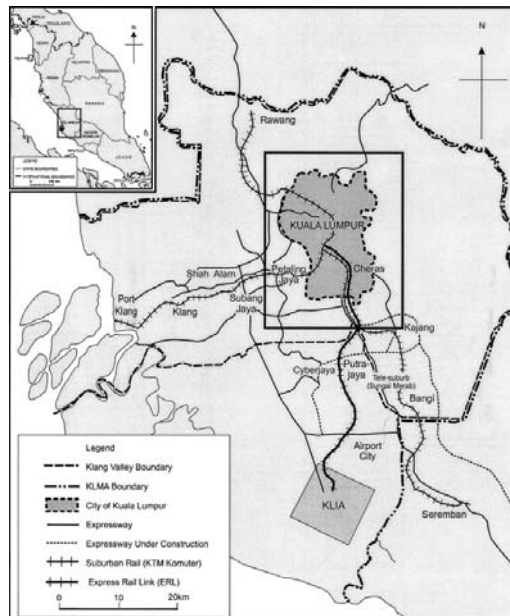
Fig. 1. : The location of Kuala Lumpur, Malaysia
Source: Bunnell (2002)

Kuala Lumpur reflects a similar trend in urban park evolution to the United Kingdom, due to its inheritance of the colonial landscape since the end of the 19th century. Kuala Lumpur (3° 8' N, 101° 41' E) the capital city of Malaysia is located within the Klang Valley, on the west coast of the Peninsula Malaysia (see Fig. 1). The city covers an area of 243 km^2, with a total population of 1.6 million as of the 2012 census report. The city has an equatorial climate characterized by hot and humid weather throughout the year, a temperature range of 21°-31° C, and 80% humidity level. There are two seasons, wet and dry. During the wet season, from April to September, Kuala Lumpur receives very high levels of precipitation, affected by the southwest monsoon, while the dry season extends from September to March.

Being a former British colony, the colonial landscape of "Picturesque and Gardenesque" style has a significant influence on design and management of urban parks in Kuala Lumpur, with a strong emphasis on beautification, and requiring active maintenance through horticultural practice. This also opens up potential for urban park management in Kuala Lumpur to adopt the ecologically sustainable landscape practices currently implemented in the UK. This is driven by the fact that Kuala Lumpur has been experiencing a similar period of rapid urban growth in line with the development of cities in the UK in the 19th century. Even so, unlike the UK, where they have built new urban parks to address social and environmental challenges facing their cities, the contribution of ecologically sustainable landscape in Kuala Lumpur is more a case of retrofitting existing parks to address these issues. Considering the climate and cultural differences, the method of adaptation and the effects of this ecological approach on the current landscapes of the city and on urban park management may differ.

3 Social and environmental issues facing the city of Kuala Lumpur, Malaysia

Social and environmental factors such as urbanization, environmental problems and climate change that are affecting urban environments worldwide have increased the need for current landscape practice to change to become more ecologically sustainable. Similarly, the tropical city of Kuala Lumpur, Malaysia has been affected by the emergence of such issues as a consequence of its rapid development since the 1970s.

The impact of urbanization and population growth on urban environment

Urbanization and population growth have significantly increased as a result of industrial development in the city. As urbanization continues to rise, urban green spaces become fragmented and slowly disappear, causing considerable loss of green space and reducing habitat diversity in the city (Byrne et al., 2009). This problem has put pressure on the local authorities to develop more urban parks and upgrade the existing parks, to fulfil the growing demand for quality outdoor spaces and environments for the people and increase biodiversity.

Rapid urban development has contributed to degradation of environmental quality, affecting the natural resources of the city, especially water. Population growth and urbanization not only create high demand for water resources, but also reduce their availability (Okun, 2000). According to Biswas (2006), as the urban population increases, the existing water supply becomes unable to support the high demand for water resources.

The process of urbanization disturbs the natural drainage of the city, creating more impervious surfaces, which affects the natural hydrological cycle. New developments usually design their drainage to fit the existing system, which results in an increased volume of stormwater, exceeding the maximum conveyance capacity of the drains, which leads to flooding (Nascimento & Baptista, 1999). There is increasingly frequent recurrence of   ash   oods due to poor drainage systems, with the higher intensity of urban activities damaging the urban infrastructure and the surroundings.

Climate change is further exacerbating environmental problems through extreme dry weather (El Nino) and severe wet weather (La Nina) conditions. Figure 2 shows the Malaysian Meteorological Department (2009) report on El Nino events that have resulted in extremely dry weather in Peninsular Malaysia, causing droughts in 1963, 1997 and 2002 that have led to a water crisis in its cities. Meanwhile the La Nina events have also led to environmental disasters such as flood events in the urban areas as the level of precipitation has increased annually, which causes frequent flooding in Kuala Lumpur, especially during the monsoon season (Mohan, Kwok, & Wan Azli, 2010). Six major floods occurred in Kuala Lumpur between the 1980s and the late 1990s, and there were further episodes of major flooding in 2001, 2002 and 2003 (Saw, 2009). The Malaysian government has spent RM1.9 billion on flood prevention measures, adopting engineering solutions such as the Stormwater Management and Road Tunnel (SMART Tunnel) Project as part of the stormwater mitigation measures for the city (ITA-AITES, 2011). However, Kuala Lumpur still suffers frequent flood events, especially during the monsoon season.
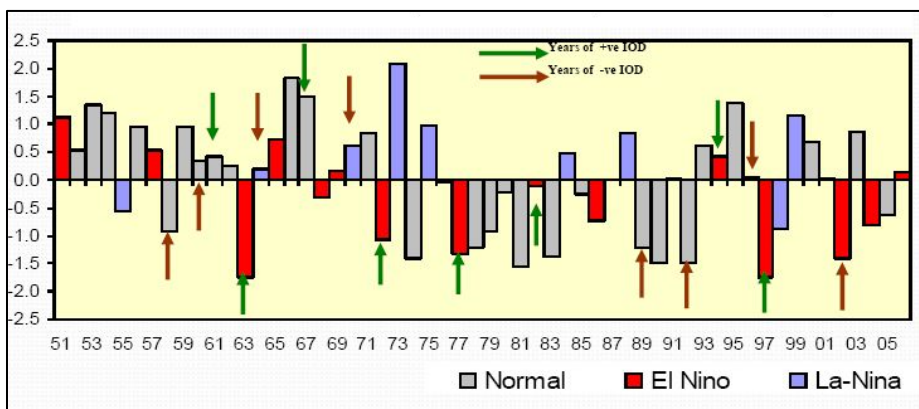
Fig. 2. : Standardized Annual Precipitation Anomaly for Peninsular Malaysia
Source: Malaysia Meteorological Department (2009)

Despite the potential of urban parks as a medium for addressing these problems in the cities of developed countries, they are not being used to address similar challenges in Kuala Lumpur, but instead, their maintenance is continuing to consume the city's resources. This has put pressure on the local authority as they are facing several issues in managing these parks, as discussed in the following section.

4 The issues in urban parks management in Kuala Lumpur, Malaysia

The need to adapt to the social and environmental challenges facing Kuala Lumpur has imposed a considerable burden on local authorities, as the current practice is still based on traditional horticultural methods, which require high maintenance. According to Justice (1986, p.178), "It is becoming more expensive to develop, manage and maintain the traditional gardens and landscapes that require specialized or constant care and attention along with watering during drought periods, weekly mowing of grass, trimming, pruning and shaping of shrubs, daily". This maintenance practice "employs a very prescriptive approach to plant care; pests and diseases are to be eliminated when found, irrespective of the damage they cause, and plants are to be watered and fertilized, irrespective of whether this is required or not" (Hitchmough and Dunnett, 2008, p. 14). This practice creates landscapes that are heavily dependent on human care to sustain their growth and incapable of adapting to the challenges of environmental change (see Fig. 3).

Fig. 3. : Intensive management of urban parks in Kuala Lumpur including blowing dried leaves, grass clippings and using potable water for irrigation
Source: Roziya Ibrahim (2012)

There is also growing pressure on government resources as a result of increasing urbanization and population growth creating greater demand for potable water, a resource which is currently used to irrigate public parks. The urban parks place a high demand on potable water for their maintenance at a time of growing demand for this resource in the city. The use of potable water for landscape irrigation is not only costly but also threatens the water supply in the city (Mohd Nor et al, 2011).

Financial restrictions are another issue affecting urban park management. Although the Malaysian government has provided a significant allocation for the creation of new parks and upgrading of old parks, no specific budget was allocated to the management and maintenance of these landscapes (Tahir, 2005). With no definite funding to maintain the urban parks, current management practices may no longer be appropriate, and an alternative, more cost-effective solution may be required. According to Justice (1986), the tropical landscape of Kuala Lumpur needs to be changed into a more ecologically sustainable landscape that would require less maintenance and should ideally be self-sustaining.

It can be summarized that in response to the environmental challenges, conventional landscape practices are no longer appropriate for managing urban parks, because of their dependence on high levels of maintenance (Cranz & Boland, 2004; Hitchmough & Dunnett, 2008) that is costly (Justice, 1986) and potentially damaging to the environment through consuming "huge amount of resources in energy for transport, irrigation and fertilizers" (Smith, Dunnett, & Clayden, 2008, p. 5). Therefore, urban park management in Kuala Lumpur could learn from the experience of developed countries in potentially adopting a more sustainable approach to addressing these issues.

5 Ecologically sustainable practice: learning from the experience of developed countries

Developed countries, such as the UK and the USA have experience in managing urban park evolution in response to social and environmental changes in their cities. They are now moving towards a more sustainable practice that aims at managing urban landscapes, including urban parks, towards ecological sustainability, and enhancing their multifunctional role of providing eco-system services beneficial to urban communities and the environment. According to Botequilha Leitão & Ahern (2002), although a similar concept was introduced by Frederick Law Olmstead in the development of New York Central Park in the 1850s, it was not fully applied until the middle of the 20th century. Ecologically sustainable practice began to be revived as a recognized practice in landscape and urban planning in the 1960s, mainly in highly urbanized cities affected by industrialization. Environmental degradation due to rapid urban development has created an urgent need for developing ecological solutions to

urban landscape management, hence promoting research into development of sustainable management approaches.
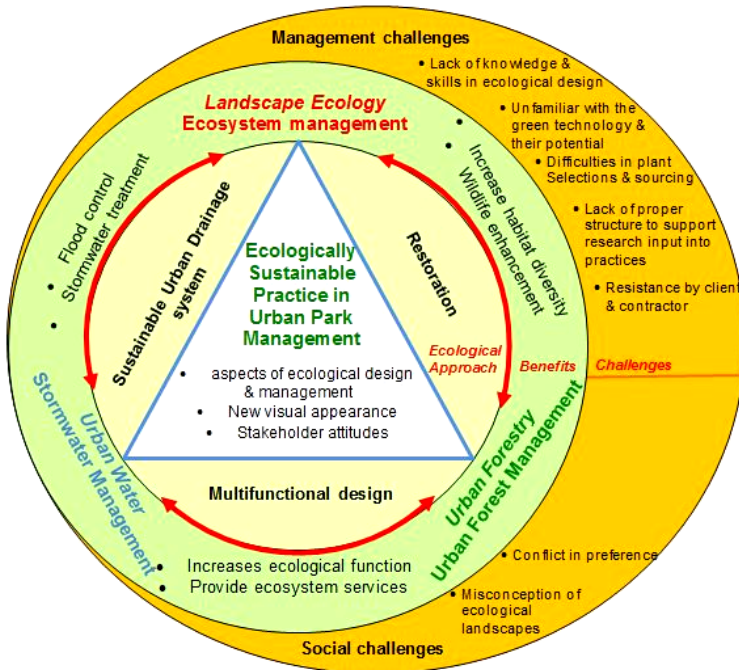
Fig. 4.: Theoretical framework for changes in urban park management towards a more ecologically sustainable practice

Figure 4 summarizes that ecological knowledge from the fields of landscape ecology, urban forestry and urban water management has informed research and practice in managing urban landscape through various ecological approaches. In the field of ecology, ecological sustainability focuses on the conservation of urban flora and fauna and their ecological processes through ecological restoration (Seabrook et al., 2011) to adapt to social and environmental challenges. The strategy for implementation of an ecological approach may vary according to the specific scale and site context. For the large urban scale, implementation could entail adaptive management (Gunderson, 2000); at a micro scale it may involve increasing the "plasticity, ecological resilience and structural diversity" (Hunter, 2011,p. 174). However, most importantly, Cranz and Boland (2004) suggest that the application of this approach at the urban park scale could potentially improve their ecological performance, which would be beneficial for the city's environmental sustainability.

Similarly, the urban forest concept could assist urban park management in conserving and managing trees and woodlands in urban parks through an arboricultural approach to mitigating the effects of urbanization (i.e. urban heat islands and pollution),

whilst at the same time offering recreational facilities for urban dwellers (Konijnendjik et al., 2006). Clark et al.'s (1997) model for urban forest sustainability placed emphasis on vegetation management practice for creation and conservation of habitat diversity that would allow ecological processes to occur naturally. Such variety of habitats makes an important contribution to increasing biodiversity and wildlife enhancement.

Recently, water management has become a necessity for improving the hydrological cycle in the city, where natural processes have been disturbed by urbanization, causing water pollution and flooding problems. Sustainable solutions for addressing stormwater in urban areas involve incorporating ecological schemes as components of the landscape, such as sustainable urban drainage systems (Kennedy et al., 2008) and artful rain gardens (Dunnett & Clayden, 2007; Echols, 2008). Both systems consist of integrated components for stormwater collection based on landscape treatment before discharge to the waterways.

The integration of these ecological approaches into urban park management could enhance their multifunctional roles by creating habitat diversity and enhancing wildlife, for the conservation of biodiversity in urban parks (Dunnett & Clayden, 2007; Lovell & Johnston, 2008; McGuckin & Brown, 1995). Besides providing public recreational spaces and facilities (Kennedy, Lewis, Sharp, & Wong, 2007), this landscape is also beneficial for the city and the urban community through "sustaining the environmental quality, resource conservation, economic development, psychological health and social well-being" (Clark et al., 1997, p.18). In addition, these approaches create a more naturalistic form of visual attraction (Dunnett & Clayden, 2007; Echols, 2008). However, the designs should consider public preference on aesthetic aspects by introducing plants that are visually pleasing (Echols, 2008). Makhzoumi (2000) suggests that an ecological approach could potentially shift the current landscape towards a more sustainable practice with higher ecological value, while at the same time retaining acceptability to the public.

The experience of developed countries in changing approaches to urban park management into a more ecologically sustainable practice could become a reference for Kuala Lumpur in adopting a similar approach that could be equally beneficial; especially the experience of the United Kingdom, because Kuala Lumpur has been adopting colonial-style landscape design and management since the end of the 19th Century. Givoni (1992) posits that there is great potential for implementing a self-sustaining landscape in tropical countries, due to high levels of precipitation, where rainwater can be utilized for irrigation of the landscape. However, the different climate conditions and cultural setting of temperate and tropical regions may create new challenges for the urban park management in Kuala Lumpur in applying such an approach to the local context. Thus, precedent studies in Singapore were chosen to fill in this gap.

Creating ecologically sustainable landscapes in tropical climates will create a different landscape character because of the different types of species and vegetation structure. It also will significantly transform the conventional landscapes into a more

naturalistic style, which are a little bit messy in their appearance. Because "different cultures have different value systems and relationship with nature" (James et al., 2009, p.69), this could generate both negative and positive perceptions amongst practitioners (Hitchmough & Woudstra, 1999) and also from the public (Nassauer, 1995a). Although ecological landscapes may provide many benefits to the people and the urban environment, Nassauer (1995) argues that they may not be acceptable to some people as they may misunderstand the naturalistic appearance as being unkempt. Their attitudes towards an ecological design and management will have a significant impact on the success or failure of this practice in tropical urban parks. Therefore, to improve cultural acceptance of ecological landscapes, Nassauer's (1995a) theory of cues to care suggested the "naturalness" form of these landscapes to be presented as they are intentionally designed for at the same time familiar to a specific culture.

Considering the aspects discussed above, this paper presents the potential benefits and challenges of implementing ecologically sustainable practices in developed countries, both in temperate and tropical climates from a management perspective, which could inform and assist the urban park management in Kuala Lumpur, Malaysia in adopting a similar approach appropriate for the local context and background. The results could offer useful knowledge and guidelines for urban parks management in Kuala Lumpur, Malaysia for appropriate development of their practice.

6 Materials and method: Review of precedent studies of urban parks which incorporate a focus on ecological design

This paper draws on analysis of interviews with the designers and managers of the selected urban parks in the UK and Singapore as exemplars of ecologically sustainable practices representing experience of different climate and cultural background, and the observed attributes of ecological design and management of each site in terms of their benefits and challenges. The relevant precedent studies were identified and selected through a review of published literature and project websites.

Precedent studies in the UK were selected because they demonstrate a greater experience of alternative approaches to managing these landscapes and innovations in their landscape design and maintenance. The examples include the Manor Fields Park (MP), Sheffield and the London Olympic Park (LOP). Meanwhile, Bishan Park-Kallang River (BP) (see Fig. 5) and Tampines Eco-Green Park in Singapore were chosen for their recent experience of implementing such an approach in tropical climates that are similar to that Kuala Lumpur, Malaysia, which also uses a similar vegetation palette.

Fig. 5.: A photos of Bishan Park-Kallang River project, Singapore (2011)

In addition, a precedent study set in the Wetland Park, Putrajaya (WP), was reviewed, representing a new appetite for ecologically sustainable practice in Malaysia, in order to validate the findings from the UK and Singapore based on the local context. The findings were supported by documents and reports relevant to the design and management of each project. Table 1 provides a brief description of the five selected precedent studies.

Table 1: Brief description of the precedent studies sites

Precedent studies	Age	Scale (ha)	Hierarchy	Location	Role	Characteristic
Manor Fields Park, Sheffield	2	25	District Park	South-east of Sheffield City	Nature conservation/stormwater management & leisure park	Hilly/ grassland/ parkland/lawn/waterbody
London Olympic Park	1	102	Regional Park	East of London	Biodiversity conservation/stormwater management Sports & Recreational park	Semi-woodland/ Parkland/Lawn/Waterbody
Bishan Park – Kallang River	26	62	District Park	Midland of Singapore	Stormwater management/ Biodiversity conservation/ & Recreational park	Semi-woodland/ Parkland/Lawn/Waterbody
Tampines Eco-Green Park	2	36.5	Regional Park	East of Singapore	Ecological park/passive recreation	Woodland/grassland/ Marshland/Lawn
Putrajaya Wetland Park	16	197	Metropolitan Park	North of Putrajaya, Malaysia	Ecological park/ Leisure & recreation	Hilly/Natural & Semi-woodland/ parkland/Lawn/wetland & lake

- **Data Analysis**

As shown in figure 6, the best management practices identified from the precedent studies were reviewed and analysed using an analytical process based on explanation-building technique (Yin, 1984) to assist in identifying potential benefits and challenges in implementing such approaches in tropical urban park management. Multiple sources of evidence, such as interviews, site observation and documentation, were used to support and validate the overall research findings through triangulation (Yin, 2009).
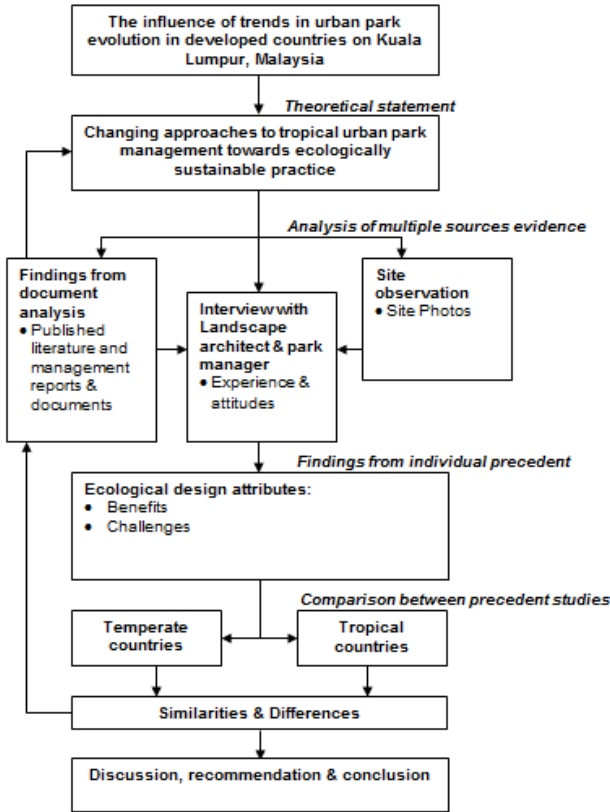
Fig. 6. : The analysis process for precedent studies based on
Yin's (2009) iterative explanation-building technique

Comparison of practices in temperate and tropical countries allowed description of similarities and differences in delivering the design and management in relation to the specific local context. The initial analysis was aimed at gaining the stakeholders' experience and perspectives on the benefits and challenges of designing and operating an ecologically sustainable landscape. The lists of benefits and challenges from the individual projects were then grouped to form several categories. Finally, these bene-fits and challenges were considered in terms of temperate and tropical countries to identify similarities and differences. This evidence was organized and presented in tables 2.2 and 2.3. The findings could inform the tropical urban park management of Kuala Lumpur in adopting a more ecologically sustainable practice appropriate for their local climate and cultural context. This paper acknowledges that there are many similar projects in other countries that are relevant to this study. However, due to restrictions of time and access to these projects, the selection of precedent studies was limited to the United Kingdom, Singapore and Malaysia.

7 Comparison of the potential benefits and challenges of ecological design and management of urban parks between temperate and tropical countries

Comparison of the benefits and challenges identified in the precedent studies highlighted certain similarities and differences between the individual projects, especially in relation to climate and cultural background. As shown in Table 2, the key findings confirmed the potential benefits of ecologically sustainable practice environmentally, socially, and economically, besides the opportunity it gives for urban park management to enhance their landscape practice.

Table 2.: Comparison of potential benefits in temperate and tropical countries

Ecological design and management benefits	Temperate countries		Tropical countries		
	MP	LOP	BP	TP	WP
1 Environmental benefits					
a. Flood control and improve stormwater quality	***	***	***	***	***
b. Increase habitat diversity for biodiversity enhancement	***	***	***	***	***
c. Enhance landscape resilience	***	**	**	***	**
d. Increase ecological performance for ecosystem services	***	***	***	***	***
2 Social benefits					
a. Enhance multifunctional use of urban park	**	***	***	**	***
b. Bring urban nature within closer proximity of residential areas	***	***	***	***	***
c. Encourage socio-ecological interaction	*	**	**	***	**
d. Foster environmental education and practice	***	***	***	***	***
e. Encourage public engagement with park management & activities	**	***	***	***	**
f. Improve public attitudes towards ecological sustainability	***	***	***	***	**
g. Promote eco-tourism		***			***
3 Economic benefits					
a. Reduce design and maintenance cost	***	***	*	***	**
b. Increase the value of surrounding land and properties	***	***	***		***
c. Generate multiple sources of funding for urban park	***		***		
4 Management benefits					
a. Enhance multidisciplinary collaboration between public and private agencies	**	***	***	*	***
b. Introduce innovative ideas to design and management	***	***	***	***	**
c. Foster research and experimental inputs into practice	***	***	***	***	**
d. Acquire and improve ecological knowledge and skills	***	***	***	***	**
e. Encourage smart partnership in park management	***	***	***		
f. Promote adaptive design and management	***	***	**	***	**
g. Reduce maintenance intensity	***	**	**	***	**

Legend: the level of benefit - Minimal* Moderate** Optimal***

- **Potential benefits**

Environmental benefits. Overall, it can be summarized that despite the specific goals and objectives of the various projects, all the studies recognized similar environmental benefits of ecological design. However, the level of benefits varied between the projects. For example, in all the projects, incorporating a sustainable urban drainage system (SuDS) into the design of parks has been very successful in flood control and

improvement of stormwater quality. Because of the integration of landscape elements with the SuDS, not only is stormwater collected, restored and managed sustainably, but also habitat diversity is created within these different elements, which helps to increase biodiversity and wildlife enhancement by providing sources of food and shelter. The multifunctional framework of the design, therefore, enhances provision of ecosystem services to the city. *"We have created a very simple yet robust system to manage stormwater on-site, improve biodiversity, reduce maintenance intensity, especially on mowing" (MP1).*

However, the different strategies for implementation of ecological design in each project reflected different levels of environmental benefit. For example the slow incremental approach of ecological design adopted in the Manor Fields Park, Sheffield and the Tampines Eco-Green Park, Singapore has been very effective in promoting a self-regenerating landscape . This is achieved because the park management adopt a vegetation strategy that places emphasis on maintaining most of the existing vegetation through natural succession (70%) and applying only minimal planting (10-30%) in specific areas, such as at an entrance. This natural process allows the ecological landscape to become more resilient to site modification.

Social benefits. In terms of the social benefits, ecological designs have the advantage of being multifunctional and integrated, thus enhancing and expanding the role of urban parks within and beyond the park boundary. By integrating the park with other green spaces through green connectors, the size of urban green spaces is increased; thus urban nature is potentially brought within closer proximinity of the residential areas. This creates opportunity for the public to experience and interact with nature and wildlife. However, except for in Tampines Eco-Green Park, this impact is still very minimal in the other parks. It is also interesting to see how each of the studied parks has used ecological design elements to create an interactive showcase, thereby providing a significant contribution in fostering environmental education and practices among the public.

Over time, the precedent study sites have successfully encouraged the public and community to engage with the park's programmes and management activities. For example, the public are involved in *"setting up the butterfly garden"(TP1);* and schools are *"very interested in being part of the park, learning more about the park (TP1)* and to participate in *"the planting programme"(WP3).* Likewise, the park also *"engaged privates companies to make contributions like picking litter, and helping to weed certain plants out"(BP2).*

Through their involvement in park management and activities, the public's attitude gradually improved and they started to give positive responses and show acceptance of the ecological park. *"At the initial stage the public considered this landscape as weedy, not maintained, and unkempt, but they slowly began to accept this approach (BP1).* And eventually, *"the approach helps them take ownership of the site"(MP2). "The local people become the park keeper by patrolling the site and keeping the river clean (BP1).*

In the city context, a larger scale of urban parks contributes to promoting eco-tourism, such as in the London Olympic Park and the Wetland Park, Putrajaya. This gives urban planners more scope to deploy appropriate physical and financial resources for delivering a high impact ecological design to attract local and international tourists.

Economic benefits. The economic benefits are distinct among the five precedent studies. Although most of the parks benefit from the reduction in landscape design and maintenance cost, the Bishan Park-Kallang River and the Wetland Park, Putrajaya show a lower decrement than the other parks. This is because the green innovations introduced, such as cleansing biotopes, soil bio-engineering techniques and the wetland system, still require constant maintenance and specialist inputs. Except for Tampines Eco-Green, all the other parks have contributed to an increase in the value of the surrounding land and properties. Manor Fields Park is one of the best exemplars of how ecological design could generate different sources of funding to support urban park management. For example, the establishment of a social enterprise helps generate additional income to support maintenance funding through various landscape services. The integration of SuDS within the park has also attracted additional funding from the drainage authority for the design and management of the park. A similar practice is also in evidence in the Bishan Park-Kallang River, Singapore through joint funding and management of the urban park and its waterways.

Management benefits. The urban park management also gains multiple benefits from ecologically sustainable practice. Most importantly, it enhances multidisciplinary collaboration between local authorities and other related agencies, bringing in local and international expertise from a range of backgrounds. However, the level of collaboration was found to differ among the studied projects. For example, the Tampines Eco-Green Park and the Manor Fields Park have a simpler organizational structure as compared to the London Olympic Park, the Bishan Park-Kallang River, Singapore and Putrajaya Wetland Park, where the projects' goals, scale and designs are more complex.

All the precedent projects recognized that collaborative efforts open up more scope for incorporation of innovative ideas and technologies in ecological design, which fosters research and turning experimental inputs into practice, especially in terms of adapting ideas and technologies from other countries into the local context. Ecological sustainable practice also helps the urban park management to acquire and improve their knowledge and skills in ecological design. *"So you begin to learn how to manage the site when it's a bit of an unknown, with such a difficult area, you just don't know what you are going to convert, it is slowly feeling your way in, growing your management knowledge and skills as the park develops (MP1).* By learning through practice, the urban park management began to promote design and management approach that is not only adaptive to environmental and social changes but *"flexible to political and economic change" (MP2).*

- **Potential challenges**

The precedent studies also revealed that the implementation of ecological sustainable practice in tropical countries created some similar challenges to practice in temperate climates (see Table. 3). However, some new challenges also emerged, which could be relevant when considering implementing such an approach to urban park management in Kuala Lumpur, Malaysia.

Table 3.: Comparison of potential challenges in temperate and tropical region

Ecological design and management challenges		Temperate countries		Tropical countries		
		MP	LOP	BP	TP	WP
1	**Environmental challenges**					
	a. Conflict between socio-ecological needs	***	***	***	***	***
	b. Disadvantage of green innovation strategy			***		**
2	**Social challenges**					
	a. Negative perception on the appearance of ecological landscape	**	**	**	**	***
	b. Poor perception on safety	**			**	***
	c. Slow acceptance of ecological landscape	***		***	***	***
3	**Economic challenges**					
	a. To define and justify the cost for sustainability	***	***	***	***	***
	b. To sustain the socio-ecological balance requires more investment	***	***	***	*	***
4	**Management challenges**					
	a. Difficult to gain mutual agreement on ecological design and management	**	**	**	***	**
	b. Difficult to change the maintenance culture among the workforce		*	**	***	
	c. Difficulties in selecting and sourcing the appropriate vegetation	*	**	**	**	*
	d. Ecological design requires elaborate maintenance	*	**	**	**	**

Legend: The level of challenge - Low* Moderate** High ***

Environmental challenges. Conflict between socio-ecological needs is a major challenge in implementing ecological design. Ecological design aims for multifunctional use of urban parks, but to break down the boundaries between these requirements is not an easy task because there is a tendency for people to avoid wildlife space and vice versa. *"When you put in a pond, it became an amenity focus, which thus limited ecological value" (MP2).* This is because the amenity zones are *"human traffic areas where a lot of wildlife will not tend to visit, especially when there is high intensity usage. Wildlife usually favoured the Upper Thompson site because some of the trees are thicker and dense. (BP2).* Fulfilling the different projects' goals and objectives created other challenges to delivering green innovative strategy at each of the precedent sites. In the UK, for example, the main goal of an ecological design would be to address the issues of limited funding for managing urban parks; thus, the management strategy may focus more on a *"light touch approach" (MP2, LOP1),* with lower resource input, especially labour. However, in Singapore and Malaysia, the main goals are to enhance the city's image (WP2) and develop environmental education and awareness (BP1) where cost is not such an issue. Besides requiring specialist input,

poor strategy in implementing these innovations consequently entails high resource and energy consumption, which is not environmentally friendly.

Social challenges. Public perception on ecological design is an important issue for both temperate and tropical countries. There are always misconceptions among the public regarding the appearance of ecological designs. *"At the initial stage, the public considered this landscape as weedy, not maintained, and unkempt" (BP1).* This is because people don't often appreciate this landscape type as something that you should have in an urban park. The public also fail to recognize the naturalistic style for safety reasons. The Manor Fields Park and Tampines Eco-Green Park are facing similar challenges in this respect; however, this is more obvious in the Wetland Park, Putrajaya. *"People are not very keen to go there because they feel unsafe, except those who are actually doing research and those who like adventure" (WP2).* Because of this perception, the public have been slow to accept ecological landscapes. *"It takes time to convince the public about this approach…sometimes you have to frame the naturalistic landscape a little bit more, perhaps make the people a little bit comfortable with it" (MP1).* This challenges the urban park management to deliver such an approach in a form that is acceptable by the public. However, the London Olympic Park has successfully gained public acceptance because *"it is a familiar vocabulary, they were very familiar with woodlands and meadows. They are not familiar with bright colour, but people are very drawn to colour, they are drawn to that kind of aesthetic" (LOP1).*

The economic challenges. Defining and justifying the cost of sustainability varies between temperate and tropical countries. The temperate countries define the cost of sustainability as using low resources and labour inputs in delivering ecological design, which helps reduce their landscape design and maintenance costs. However, in the tropical countries, delivering an ecological design entailed big investment, especially in green technologies such as the wetland system, soil bio-engineering technique and the cleansing biotope, which means that the maintenance costs could be similar to those in a conventional park until the landscape is established. Furthermore, to sustain the socio-ecological balance of the parks in the future requires more investment; for example, to improve people's experience of the parks through programmes and activities, and at the same time manages their ecological elements to provide other ecosystem services.

The management challenge. The issues described above created major challenges for urban park management in delivering ecological designs. These projects involved multidisciplinary collaboration and it is always a challenge to convince the relevant stakeholders of the merits of an ecological design, as well as to fulfil their different needs and requirement from the project. This issue is most critical when it involves a complex organizational structure, as in the cases of the London Olympic Park and the Putrajaya Wetland. It is also difficult to change the maintenance culture towards ecologically sustainable practice when the management team is used to the conventional approach, especially in a project with a higher ecological aim, such as the Tampines

Eco-Green Park. Because delivery of ecological design needs to be balanced with the public's requirements, plant selections and sourcing is always problematic, especially when the project places emphasis on local species and sources of these are very limited. Some ecological design projects include appropriate exotic species with high ecological value and of more attractive appearance, but these need to be imported from other countries. However, in the UK, The London Olympic Park and Manor Fields Park have started collaborating with the local universities in experimenting with exotic species and planting techniques appropriate to their local climate, such as sowing techniques. Although ecological design introduces techniques that help to reduce the necessary vegetation management, some of these techniques require elaborate maintenance. *"The maintenance specification explains not just how you do it, but also why you do it" (BP1)*. Therefore, constant supervision of the maintenance work is required, besides proper training to customize the knowledge and skills of the workforce.

8 Discussion

Comparison between these urban parks reveals that there are both similarities and differences in terms of the benefits and challenges in the delivery of ecologically sustainable practice by the respective local authorities, subject to their different roles in providing ecosystem services for the city, which vary according to the park's scale and its location within the city.

Socially, all the precedent studies reflected similar benefits accruing from multifunctional designs that not only provide recreational space for the public, but also help instil an awareness of environmental issues. This growing awareness may help to improve the public's attitudes towards ecological landscapes by developing a sense of ownership and encouraging their active involvement with the park management and activities. In terms of ecological aspects, although respect for the existence of urban nature and wildlife among the public is in evidence in all the precedent sites, such socio-ecological interaction is still considered low, which imposes a similar challenge to the management of all the parks; and more design and management input is required to balance the two entities. In terms of economic benefits, the majority of the parks have successfully increased the value of the surrounding land and properties.

Despite having similar benefits and challenges, it is observed that parks of different scale and location also offer a different level of ecological benefits and challenges, which can be described in terms of the degree of ecologically informed design and management pertaining to each specific park. It is observed that the smaller the park, the greater the adoption of ecological design, whereas provision of recreational facilities may be reduced and vice versa. Lack of knowledge of ecological design has become a challenge to local authorities in presenting this idea to the public and has resulted in poor public and management perception of the safety and appearance of ecological landscapes. However, the response towards such a landscape varies according to social and cultural backgrounds.

All the above criteria are also influenced by the complexity of the landscape management structure and strategies for each park. The precedent studies have shown a good exemplar of how the willingness of urban park management can change conventional practice to a more sustainable practice in responding to emerging social and environmental issues and challenges in the city. Even so, variations in scale and location have led to the emergence of distinct sets of priorities and strategies among the respective local authorities in managing these public spaces towards ecological sustainability. Larger scale parks, located generally in the centre of the city, potentially receive greater financial support and management resources from the local authority because of their greater contribution towards a wider city green infrastructure. Meanwhile, smaller scale parks, located in outlying areas, usually suffer from budget constraints; thus, financial allocation and operational and maintenance intensity for such parks are much lower than for the larger scale parks. This suggests that implementation of ecological design and management does not mean merely minimum spending. The overall cost of ecological design depends on the management's understanding of this approach; and their strategy in applying appropriate technology for the specific context of the city.

9   Conclusion

The paper demonstrates the potential benefits and challenges in implementing ecologically sustainable practice in urban parks. Besides showing similarities and differences between temperate and tropical countries in delivery of ecological design outcomes, it identifies unique strengths and weaknesses in delivering such an approach in a specific local context. These outcomes contribute to the potential development of alternative design and management guidelines for managing tropical urban parks, which the local authority responsible for managing urban parks in Kuala Lumpur could then evaluate further in terms of the practicality and adaptability of such an approach to their local context. The overall findings could inform and raise awareness among tropical urban park management in Kuala Lumpur, towards delivering a more ecologically sustainable practice.

10   Acknowledgement

References

Antrop, M. (2005). Why landscapes of the past are important for the future. *Landscape and Urban Planning, 70*(1-2).

Biswas, A. K. (2006). Water management for major urban centres. *International Journal of Water Resources Development, 22*(2), 183-197.

Botequilha Leitão, A., & Ahern, J. (2002). Applying landscape ecological concepts and metrics in sustainable landscape planning. *Landscape and Urban Planning, 59*(2), 65-93.

Byrne, J., Wolch, J., & Zhang, J. (2009). Planning for environmental justice in an urban national park. *Journal of Environmental Planning and Management, 52*(3), 365-392.

Calkins, M. (2005). Strategy use and challenges of ecological design in landscape architecture. *Landscape and Urban Planning, 73*(1), 29-48.

Clark, J. R., Matheny, N. P., Cross, G., & Wake, V. (1997). A model of urban forest sustainability. *Journal of Arboriculture, 23*, 17-30.

Conway, H. (2000). Parks and people: the social functions. In J. Woudstra & K. Fieldhouse (Eds.), *The regeneration of public parks* (pp. 9-20). London: Taylor & Francis Group.

Cranz, G., & Boland, M. (2004). Defining the sustainable park: A fifth model for urban parks. *Landscape Journal, 23*(2), 19.

Dunnett, N., & Clayden, A. (2007). *Rain gardens : managing water sustainably in the garden and designed landscape*. Portland, Or.: Timber.

Echols, S. (2008). Artful rainwater design in the urban landscape. *Journal of Green Building, 2*(4), 19.

Givoni, B. (1992). Climatic aspects of urban design in tropical regions. *Atmospheric Environment. Part B. Urban Atmosphere, 26*(3), 397-406.

Gobster, P., Nassauer, J., Daniel, T., & Fry, G. (2007). The shared landscape: what does aesthetics have to do with ecology? *Landscape Ecology, 22*(7), 959-972.

Gunderson, L. H. (2000). Ecological resilience - in theory and application. *Annu. Rev. Ecol. Syst, 31*, 425-439.

Hitchmough, J., & Dunnett, N. (2008). Introduction to Naturalistic Planting in Urban Landscapes. In N. Dunnett & J. Hitchmough (Eds.), *The Dynamic Landscape* (pp. 1). New York, USA.: Taylor & Francis.

Hitchmough, J., & Woudstra, J. (1999). The ecology of exotic herbaceous perennials grown in managed, native grassy vegetation in urban landscapes. *Landscape and Urban Planning, 45*(2–3), 107-121.

Hunter, M. C. (2011). Using Ecological Theory to Guide Urban Planting Design An adaptation strategy for climate change. *Landscape Journal, 30*(2), 173-193.

ITA-AITES. (2011). SMART-Malaysia. from http://uww.ita-aites.org/index.php?id=223

Jordan, H. (1994). Public Parks, 1885-1914. *Garden History, 22*(1), 85-113.

Justice, C. L. (1986). The concept of the urban forest as applied to kuala lumpur, malaysia. *journal of arboriculture, 12*(7), 4.

Kennedy, S., Lewis, L., Sharp, E., & Wong, S. (2007). *Sustainable Urban Drainage System (SUDS) - More than a drainage solution?* Paper presented at the NOVATECH 2007 - 6th International Conference on Sustainable Techniques and Strategies in Urban Water Management, Lyon, France.

Konijnendijk, C. C., Ricard, R. M., Kenney, A., & Randrup, T. B. (2006). Defining urban forestry – A comparative perspective of North America and Europe. *Urban Forestry & Urban Greening, 4*(3–4), 93-103.

Lawrence, H. W. (2008). *City Trees: A Historical Geography from the Renaissance Through the Nineteenth Century*: University of Virginia Press.

Lovell, S. T., & Johnston, D. M. (2008). Creating multifunctional landscapes: how can the field of ecology inform the design of the landscape? *Frontiers in Ecology and the Environment, 7*(4), 212-220.

Makhzoumi, J. M. (2000). Landscape ecology as a foundation for landscape architecture: Application in Malta. *Landscape and Urban Planning, 50*(1-3), 167-177.

Malaysian Meteorological Department. (2009). *Climate Change Scenario for Malaysia 2001-2099*.

McGuckin, C. P., & Brown, R. D. (1995). A landscape ecological model for wildlife enhancement of stormwater management practices in urban greenways. *Landscape and Urban Planning, 33*(1–3), 227-246.

Mohan, K. S., Kwok, L. L., & Wan Azli, W. H. (2010). *Extreme Climate Change Scenarios over Malaysia for 2001-2099*. Petaling Jaya, Selangor: Malaysian Meteorological Department (MMD), Ministry of Science, Technology and Innovation (MOSTI).

Mohd Nor et al. (2011). *Current Trends in Integrated Urban Stormwater Management Practice in the Tropical Region Malaysia*. Paper presented at the 12th international Conference on Urban Drainage. Retrieved from http://web.sbe.hw.ac.uk/staffprofiles/bdgsa/temp/12th%20ICUD/PDF/PAP006155.pdf

Nascimento, N., & Baptista, M. (1999). *Stormwater management problems in a tropical city - the Belo Horizonte case study*. Paper presented at the Impacts of Urban Growth on Surface Water and Groundwater Quality, Birmingham, UK.

Nassauer, J. (1995). Culture and changing landscape structure. *Landscape Ecology, 10*(4), 229-237.

Nassauer, J. I. (1995). Messy ecosystems, orderly frames. *landscape journal*(15), 9.

Okun, D. A. (2000). Water Reclamation and Unrestricted Nonpotable Reuse:A New Tool in Urban Water Management. *Annual Review of Public Health, 21*, 23.

Saw, H. S. (2009). *A Case Study of Mitigating Flood in City Centre of Kuala Lumpur*. Paper presented at the Innovative Strategies towards Flood Resilient Cities in Asia-Pacific. Retrieved from http://www.unescap.org/idd/events/2009_EGM-DRR/Malaysia-Hinseang-Saw-STORMWATER-MANAGEMENT-AND-ROAD-TUNNEL_final.pdf

Seabrook, L., McAlpine, C. A., & Bowen, M. E. (2011). Restore, repair or reinvent: Options for sustainable landscapes in a changing climate. *Landscape and Urban Planning, 100*(4), 407-410.

Smith, C., Dunnett, N., & Clayden, A. (2008). *Residential Landscape Sustainability: A Checklist Tool*: Wiley.

Tahir, O. M. (2005). *Urban landscape management in Malaysia : in search of a sustainable management system*. Unpublished Ph. D., University of Newcastle upon Tyne, University of Newcastle upon Tyne, Newcastle upon Tyne.

Yin, R. K. (1984). *Case study research : design and methods*. Beverly Hills ; London: Sage.

Yin, R. K. (2009). *Case study research : design and methods* (4th ed.). Los Angeles, [Calif.] ; London: SAGE.

Zipperer, W. C., & Zipperer, C. E. (1992). Vegetation responses to changes in design and management of an urban park. *Landscape and Urban Planning, 22*(1), 1-10.

Relationship between Self-Esteem and Health-Related Behaviour

Petr Klimeš[1]

[1] Department of Psychology, Faculty of Arts, Charles University in Prague,
`klimes.petr@gmail.com`

Abstract. This paper deals with relationship between self-esteem and health-related behaviour. Concerning that previous studies used mainly correlation of self-esteem with few and separated variables of health-compromising behaviour, we decided to utilize more complex approach. We included in our study also variables of health-supportive behaviour and used factor and cluster analysis to obtain several types of lifestyle instead of working with separate variables. Besides of Rosenberg Self-esteem Scale we also used Tafarodi and Swann Self-liking/Self-competence Scale to identify inner structure of self-esteem. The study was conducted on population of 376 participants, mainly college students. The results showed one type of lifestyle that differed in self-esteem from the others and some weak correlations between self-esteem and factors of health-supportive behaviour. We were unable to confirm either any statistically significant relation between self-esteem and factors of health-compromising behaviour or relation with single variables of health-compromising behaviour.

Keywords: *self-esteem, health-related behaviour, health-compromising behaviour, health-supportive behaviour*

1   Introduction

Self-esteem as an emotional attitude to oneself and as a part of self-concept is one of the most central concepts of psyche. The enduring, continuous and generalized construct is usually called trait self-esteem. As a generalized and continuous representation of an attitude to oneself, the self-esteem suggests the importance in daily life (Blatný, 2001; Blatný et al., 2010). It is the reason why we choose this construct as basis for our research.

The construct we decided to investigate in connection with self-esteem was health-related behaviour. The reason for choosing this construct lies in critical situation for health-compromising behaviour in population of Czech Republic (Csémy et al., 2006; ÚZIS, 2004; ÚZIS, 2006; Hibell et al., 2012) and in the expected connection with self-esteem as important determinant of this behaviour. The health-related behaviour comprises of health-compromising behaviour, which includes e.g. smok-

ing, alcohol consumption, drug abuse and health-supportive behaviour, which includes e.g. healthy diet, exercising and other physical activity.

There was number of researches conducted on relationship between self-esteem and health-related behaviour. Somewhat problematic is fact, that these researches were conducted mainly on single or only several variables of health-related behaviour and were almost exclusively focused on variables of health-compromising behaviour. The relationship with health-supportive behaviour was mostly omitted. Some of these studies found connection between self-esteem and particular variables of health-compromising behaviour (Abernathy et al., 1995; DeHart et al., 2008), but mostly there was not found any relationship (Glendinning & Inglis, 1999; Kounenou, 2010; Baumeister et al., 2003).

On the basis of these results, we decided to investigate the relationship between self-esteem and health-related behaviour in a more complex way. Apart from variables of health-compromising behaviour, we included in our research also variables of health-supportive behaviour. We also choose more complex way of data processing. At first we used factor analysis on variables of health-related behaviour to obtain underlying factors, and then we used cluster analysis to get basic types of lifestyle. Our assumption was that if there is not apparent relationship between self-esteem and health-related behaviour when the variables of health-related behaviour are investigated separately, there could be found some relationship when these variables are taken as complex type of lifestyle.

In our research we tested following hypotheses:

1. Level of self-esteem varies in different lifestyles.
2. The relationship between self-esteem and lifestyle types differs for sub-scales of SLCS-R.
3. The average number of days when the individuals were ill will differ for each type of lifestyle.

2   Method

Participants

Most of the participants were contacted via Masaryk University in Brno and Charles University of Prague. Further individuals were related to the author. The total comes to 376 participants, 98 men and 278 women. Mean age of respondents was 29.3 years with standard deviation 11.5, with 66% of participants aged between 20 and 30 years; 65% of respondents were college students.

Measures

Trait self-esteem
For self-esteem measurement we utilized the following two scales: the Rosenberg 10-item self-esteem scale (SES) as a common measure of global self-esteem (Blascovich

& Tomaka, 1991), and Tafarodi and Swann (1995) 16-item Self-Liking/Self-Competence Scale Revised (SLCS-R) for revealing the inner structure of global self-esteem. For Rosenberg scale, 4-point Likert scale (1 = disagree, 4 = agree) was used, so the point range was between 10 (lowest) and 40 (highest). The mean of the scale is 25 points. The Self-Liking/Self-Competence Scale consists of two inner scales: self-liking and self-competence. Each of them comprises of 8 items. In this case participants responded using 5-point Likert scale (1 = strongly disagree, 5 = strongly agree) where the minimum is 16 points and the maximum 80. The mean of the scale is 48 points. Negative items in both scales were reverse-scored, higher scores thus indicating higher self-esteem.

Health-related behaviour questionnaire

This questionnaire was composed of items concerning health-compromising and health-supportive behaviour. The questions covering health-compromising behaviour dealt with smoking (currently and in the past), alcohol consumption (mean consumption, mean frequency, frequency in the last month, drinking five or more units of alcohol on a single occasion, intoxication in the last month and last year), drug abuse (any past use of illicit substances yes/no, marijuana consumption in last year and last month, and consumption of the following substances in last month, year and earlier: marijuana / hashish, ecstasy, sedatives, methamphetamine / amphetamines, cocaine / crack, heroin / opiates, LSD, psilocybe mushrooms / natural hallucinogens).

For health-supportive behaviour, the items included preventative usage of supplements and vitamins, regular medical check-ups, number of hours engaged in sports per week, frequency of other physical activities (walking, cycling, etc.), regular drinking regime, sleeping at least 7 hours regularly, rate of eating sweets between main meals, how healthy do the respondents think their diet is, rate of drinking coffee or strong tea). The items were mostly stated on 5 point scale regarding frequency of particular behaviour, with 1 = never and 5 = very often.

Furthermore, we surveyed the participants' rate of sickness. We used the following criteria to capture this variable: number of days of sickness in the last year, number of days when the respondents had to stay in bed, number of days they spent in a hospital, and chronic illness and somatic problems (headaches, sickness, etc.).

Data processing method

Due to the number of variables and reasons mentioned earlier, we decided to process data into "lifestyles", using factor and cluster analysis.

The factor analysis utilized 21 items gathered from a health-related behaviour questionnaire. On the basis of a scree graph, we decided to extract 4 factors from the original items. We used rotation Verimax normalized. These four factors explain 48.5% of variance in the original data. The first factor is saturated with items related to alcohol consumption. The second factor is saturated with items regarding the regularity of lifestyle and meeting basic needs (regular sleep, drinking regime and diet). The third factor consists of items concerning smoking and drug use. The last factor is saturated with items related to physical activity and healthy diet. The variables satu-

rate mostly only one factor each, saturation of other factors is usually low. On the basis of these results, we labelled the factors thus:

1. alcohol consumption
2. regularity of lifestyle
3. smoking and using drugs
4. physical activity and healthy diet

As the next step, we subjected the gathered data to cluster analysis in order to obtain the desired lifestyle types. Having considered several numbers of clusters we picked a four-cluster model, as it turned out to be the best one for interpretation. A brief characterization of main attributes of each cluster follows:

1. cluster: "non-regulars" (87 individuals)
- low alcohol consumption
- low regular satisfaction of basic needs
- average smoking and marijuana use
- average physical activity and diet healthiness

2. cluster: "non-exercising smokers" (103 individuals)
- low alcohol consumption
- relatively high regular satisfaction of basic needs
- high smoking and marijuana use
- low physical activity and diet healthiness

3. cluster: "alcohol consumers" (75 individuals)
- high alcohol consumption
- average regular satisfaction of basic needs
- average smoking and marijuana use
- average physical activity and diet healthiness

4. cluster: "healthily behaving" (111 individuals)
- low alcohol consumption
- high regular satisfaction of basic needs
- low smoking and marijuana use
- high physical activity and diet healthiness

Based on dominant characteristics we labelled the clusters as "non-regulars", "non-exercising smokers", "alcohol consumers", "healthily behaving". The distribution of individuals across all clusters is relatively even. None of them has an extremely low or extremely high number of respondents. One cluster ("healthily behaving") is characterized by a high rate of health-supportive behaviour and low rate of health-compromising behaviour. The other three clusters have high values in characteristics of corresponding factors of health-compromising behaviour, making the interpretation of the model quite clear.

3   Results

In this section we present the results of our study in two main parts. First part brings some descriptive statistics i) for basic demographic characteristics of participants regarding their lifestyles, ii) for self-esteem tests, iii) for variables of health-compromising and health-supportive behaviour and iv) for variables regarding the state of health. In second part, three hypothesis stated earlier are tested.

Differences between types of lifestyle related to socio-demographic data

Table 1 displays the differences between men and women, regarding lifestyle types they belong to.

Table 1: Distribution of participants regarding sex and types of lifestyles

sex		type of lifestyle				total
		non-regulars	non-exercising	alcohol consumers	healthily behaving	
men	number	26	18	26	28	98
	% in type	29.9	17.5	34.7	25.2	26.0
	% in sample	26.5	18.4	26.4	28.6	100.0
women	number	61	85	49	83	278
	% in type	70.1	82.5	65.3	74.8	74.0
	% in sample	21.9	30.6	17.6	29.9	100.0
total	number	87	103	75	111	376
	% in sample	23.1	27.4	20.0	29.5	100.0

The table shows that in all clusters, there are more women than men; this is due to the much higher representation of women in our sample. The percentage proportions in the sample are more informative. Here we can see a larger proportion of women in the "non-exercising smokers" and "marijuana users" clusters and a larger proportion of men in the "alcohol consumers" cluster. Smaller, but still noticeable is the ratio of men to women in the "non-regulars" cluster, which has slightly more men than women. While the differences in representation of men and women in relation to lifestyle types are noticeable, they are not statistically significant, though not far from that boundary (p=0.057).

Results of self-esteem tests

Table 2 displays the results of Rosenberg Self-esteem scale (SES) and Tafarodi and Swann Self-liking/self-competence scale (SLCS-R) and its sub-scales (SL and SC).

Table 2: Results of self-esteem tests

	Mean	Minimum	Maximum	SD
SES	31.95	12.0	40.0	5.24
SLCS-R	53.76	16.0	78.0	10.45
SL	29.04	8.0	40.0	6.18
SC	24.72	8.0	38.0	5.27

The mean value of self-esteem in SES test was 31.95 points with sd = 5.24. In SLCS-R, it was 53.76 points and sd = 10.45. In self-liking subscale of SLCS-R, there was slightly higher mean score (29.04) than in self-competence subscale (24.72). The distribution of values of SLCS-R was also much more even than distribution of SES values. It is given by limited positive spectrum of questions in SES (SES usually asks: "Do you feel good enough about yourself?" while SLCS-R usually asks: "Do you feel like the best person?"). We didn't find any differences in self-esteem regarding sex, age, and student/non-student status. Close to statistical significance were the differences in self-esteem considering religious belief (F = 3.13, p = 0.078); with higher self-esteem in non-religious individuals.

Results for health-compromising behaviour variables

Smoking
The majority of the respondents (80%) don't smoke, 7% smoke occasionally, 3.7% smoke daily (1–5 cigarettes), 4.8% smoke 6–20 cigarettes daily and only one respondent reported smoking more than 20 cigarettes per day. An interesting observation is that 4% of respondents don't smoke cigarettes but pipes or cigars.

Alcohol consumption
There are only 5% complete abstainers; 24% reported consuming alcohol once or less per month, 44% consume alcohol 2–4 times per month, 21% drink 2-3 times per week, and 4% drink 4 times or more per week. The results on binge drinking are: 7% of respondents consume 5 or more alcohol units during one occasion at least once per week, 12% at least once in a month, 41% less than once in a month and 40% never consume this amount of alcohol during one occasion.

Drug abuse

36% of respondents stated they sometime used a drug and 64% of respondents stated they never used a drug. As expected, marijuana was the most used of the covered drugs. 10% of respondents used it in the past year and 5% of respondents last month. The only other substances worth mentioning are natural hallucinogens – 7% respondents in total have ever used them. It is also interesting that 10 people (2.5%) had previously used methamphetamine.

Weight

We derived BMI values from items concerning weight and height. The results show that the majority of respondents (73%) fall into the "norm" category.

Results for health-supportive behaviour variables

Physical activity

This category contains two items: "number of hours spent exercising per week" and "frequency of other physical activity".

Results show that on average the respondents spent 3 hours a week exercising (sd = 3.56), with minimum 0 and maximum 30 hours per week. For item „frequency of other physical activity", the middle value was the most frequent one (33%), with steady descent to extreme values.

Regular basic needs satisfaction

This variable consists of "eating regularly" (3–5x per day), "sufficient amount of sleep" (at least 7 hours per day) and "sufficient hydration" (at least 2 litres per day).

All three items show a steady increase of the number of respondents with the increase of value (few respondents in first category, more respondents in last category). The variable "sufficient amount of sleep" has noticeably more respondents in the last category (47%).

Healthy diet

This variable was measured with items "healthiness of diet" and "degree of drinking strong tea or coffee". The "healthiness of diet" had respondents evaluate how healthy or unhealthy their diet is. Most individuals responded in the middle category (48%), the second highest amount of respondents choose the fourth category (32%).

Prophylactic behaviour

This category contains variables "preventative use of drugs and vitamins" and "regular medical check-ups". Results show that 62% of respondents in variable "preventative use of drugs and vitamins" fall into first and second category (never and almost never) with following categories decreasing. Variable "regular medical examination" shows similar distribution of respondents.

Results for variables regarding state of health

The questionnaire contained three variables measuring morbidity, including "number of days of sickness in the last year" and "number of days spent in bed in the last year".

The average number of days of sickness is 11.8 (sd = 13.3) according to the results; the median of this variable was 7. For "number of days in bed" the average was 5.9 (sd = 6.6).

Hypotheses verification

We are now going to attempt to verify the tree hypotheses presented in the introduction section of this paper.

1. hypothesis: Levels of self-esteem vary in different lifestyles.
We confirmed this hypothesis by analysis of variance of SES and SLCS-R tests and subscales of SLCS-R with types of lifestyle. The basic results can be seen in the Table 3.

Table 3: Analysis of variance of self-esteem and types of lifestyle

	F	p
SES	**3.07**	**0.028**
SLCS-R	2.31	<u>0.076</u>
SL	1.19	0.314
SC	**3.09**	**0.027**

The table shows that on 5% level of significance the results for Rosenberg SES and Self-competence subscale of SLCS-R are significant. We underlined also whole SLCS-R value, because it is close to statistical significance. For the Self-liking subscale of SLCS-R, there is not statistically significant difference among the types of lifestyle, in fact the results are quite far apart from it.
In the Table 4 you can see values of self-esteem (SES) for each type of lifestyle separately.

Table 4: SES results for types of lifestyle

	mean SES	SES N	SES sd
"non-regulars"	30.48	87	5.85
"non-exercising"	32.38	103	5.00
„alcohol consumers"	32.21	75	4.75

„healthily behaving"	32.53	111	5.14
all clusters	31.95	376	5.24

The table shows interesting results. "Non-exercising smokers", "alcohol consumers" and "healthily behaving" have almost even rate of self-esteem. In our lifestyle typology, there is one type („non-regulars"), which has significantly different results in self-esteem, than the others. LSD test has confirmed statistically significant difference in self-esteem (SES) between 1. cluster (non-regulars) and 2., 3. and 4. cluster (healthily behaving). Although we didn't find out differences between all types of self-esteem we did find one cluster, which significantly differs in self-esteem from all other clusters. Thus we can say the first hypothesis has been confirmed.

2. hypothesis: The relationship between self-esteem and types of lifestyle differs for subscales of SLCS-R

Table 5: Analysis of variance of subscales of SLCS-R and types of lifestyle

subscale	F	p
SL	1.19	0.314
SC	**3.09**	**0.027**

From Table 5 it is apparent, that statistically significant difference in self-esteem is present only in Self-competence subscale of SLCS-R test. It is also interesting, that results for Self-liking subscale are from statistical significance quite distant. In the following Table 6 you can see concrete values of self-esteem for Self-competence subscale for every type of lifestyle.

Table 6: Self-competence subscale results for types of lifestyle

lifestyle	mean SC	SC N	SC sd
"non-regulars"	23.50	87	5.75
"non-exercising smokers"	24.42	103	5.05
"alcohol consumers"	25.09	75	5.54
"healthily behaving"	25.69	111	4.69
all clusters	24.72	376	5.27

The "non-regulars" cluster has the lowest value of self-esteem. This is the same result as for SES. Further, we can see, that values of self-esteem rise for "non-exercising smokers" and "alcohol consumers" with highest value for "healthily behaving". Also

in this part the results resembles outcomes of SES. The difference in self-esteem for self-competence subscale of SLCS-R applies for difference between 1. cluster (non-regulars) and 4. cluster (healthily behaving), which was confirmed with LSD test. On the basis of these results we can say there is a difference between Self-liking and Self-competence subscale of SLCS-R regarding various types of lifestyle. Thus our hypothesis number 2 has been also confirmed.

3. hypothesis: Mean number of days when the individuals were ill will differ for each type of lifestyle.

We tested a connection between number of days when the individuals were ill and different types of lifestyle. We found neither relation with number of days, when the individuals were ill ($p = 0.56$), nor number of days, when they were forced to stay in bed ($p = 0.36$).

Correlation of self-esteem and factors of health-related behaviour

Table 7 shows statistically significant results of correlation of self-esteem tests and factors of health-supportive behaviour. Nevertheless these correlations are quite weak. We did not find out any significant correlation of self-esteem and factors of health-compromising behaviour. All tests except Self-liking subscale of SLCS-R correlate with factor "regularity of daily rhythm" (which regards eating regularly, sufficient hydration and sufficient amount of sleep). With factor "physical activity and healthy diet" correlate all tests without exception. Concerning these results we can say, that respondents in our sample do not differ in level of self-esteem regarding health-compromising behaviour but differ in level of self-esteem regarding health-supportive behaviour. Further we will briefly mention some significant relations with individual items.

Table 7: Correlation of self-esteem and factors of health-related behaviour

	"alcohol consumption"	"regularity"	"smoking and marijuana using"	"physical activity and healthy diet"
SES	-0.052	**0.116**	-0.006	**0.107**
	p = 0.920	**p = 0.024**	p = 0.913	**p = 0.037**
SLCS-R	-0.047	**0.103**	0.006	**0.161**
	p = 0.367	**p = 0.047**	p = 0.910	**p = 0.002**
SL	-0.069	0.072	0.007	**0.134**
	p = 0.185	p = 0.163	p = 0.894	**p = 0.009**
SC	-0.012	**0.119**	0.003	**0.161**
	p = 0.816	**p = 0.021**	p = 0.947	**p = 0.002**

Relation of self-esteem and individual items of health-related behaviour

Under the scope of our research we verified also relation between self-esteem and items of health-related behaviour. We did not find any significant relation with items of health-compromising behaviour and we confirmed relation between self-esteem and some items of health-supportive behaviour. The statistically significant relation was confirmed with items: "number of hours per week spend exercising" (SES, $p = 0.02$), "frequency of another physical activity" (SLCS-R, SC, $F = 2.98$, $p = 0.019$), "sufficient hydration" (SES, $F = 2.43$, $p = 0.047$), "eating regularly" (SES, $F = 3.67$, $p = 0.006$).

Further we found out relation between self-esteem and health condition in these items: "number of days of illness in the last year" ($r = -0.16$, $p = 0.002$), "number of days in bed in the last year" ($r = -0.11$, $p = 0.03$) and "number of somatic complaints" (e.g. headaches, sickness, heart beating,...) ($r = -0.34$, $p = 0.000$).

4   Discussion

Regarding previous research of self-esteem and health-related behaviour (Baumeister et al., 2003), we decided to proceed in different way. Most of previous researches were focused on health-compromising behaviour and examined this area relating self-esteem with individual variables. In our research, we decided to include both items of health-compromising behaviour and items of health-supportive behaviour. We also choose different approach to data processing. Considering weak results of previous studies we decided to approach health-related behaviour in a more complex way as types of lifestyle. To realize this we processed data with factor and cluster analysis.

At first we extracted 4 factors of 21 items of health-related behaviour. The first factor "alcohol consumption" was saturated with items regarding frequency of alcohol consumption, amount of alcohol consumed and with items of intoxication. The second factor was saturated with items regarding the basic needs satisfaction (items: "eating regularly", "sufficient hydration" and "sufficient amount of sleep"). We called this factor "regularity of daily rhythm". The third factor "smoking and using marijuana" was saturated with items related to the name of the factor. The fourth factor "physical activity and healthy diet" was saturated with items regarding physical activity ("number of hours spend exercising" and "frequency of other physical activity") and healthiness of diet. Utilizing cluster analysis, we derived from these four factors four types of lifestyle. The first type was labelled "non-regulars", because individuals in this cluster differed from others in the factor "regularity of daily rhythm". Individuals in the second type were called "non-exercising smokers". These individuals differed from others in higher score in factor "smoking and using marijuana" and lower score in factor "physical activity and healthy diet". Individuals in the third type differed in higher score in factor "alcohol consumption", thus we called them "alcohol consumers". Individuals in the fourth type were called after higher scores in both factors regarding health-supportive behaviour and low scores in both factors of health-compromising behaviour, as "healthily behaving".

Further we related factors of health-related behaviour as well as types of lifestyle we got from cluster analysis with results of self-esteem tests (Rosenberg Self-Esteem Scale and Tafarodi and Swann Self-liking/Self-competence Scale Revised).

The results showed a statistically significant difference is self-esteem (SES) for type "non-regulars" compared to other three types ("non-exercising smokers", "alcohol consumers" and "healthily behaving") which had almost the same score in self-esteem. The type of "non-regulars" had significantly lower score in self-esteem than the others. For SLCS-R test we also found statistically significant difference in self-esteem among the types of lifestyle. But it concerned only Self-competence subscale. In this case, there was found a significant difference in self-esteem between "non-regulars" type and "healthily behaving" type. Contrary to our assumptions, we did not find any relationship between types of lifestyle and morbidity.

Further we correlated self-esteem scores with factors of health-related behaviour. We got statistically significant results between self-esteem and factors of health-supportive behaviour. For factor "regularity of daily rhythm" the correlation was significant for SES, SLCS-R and it's Self-competence subscale. For factor "physical activity and healthy diet" the correlation applied for SES, SLCS-R and both subscales. Nevertheless these relations were rather weak. The relation between self-esteem and health-supportive behaviour was confirmed also on the level of some individual variables. We found correlation between self-esteem and "number of hours spend exercising per week", "frequency of other physical activity" (only for SLCS-R), "sufficient hydration", "eating regularly" and "healthiness of diet" (only for Self-competence subscale of SLCS-R).

In our study we were unable to confirm either relationship between self-esteem and factors of health-compromising behaviour or any relationship between self-esteem and individual items of health-compromising behaviour. We found a relation with factors and some items of health-supportive behaviour. We also found a lower level of self-esteem in a lifestyle of "non-regulars" in comparison with three other lifestyles ("non-exercising smokers", "alcohol consumers" and "healthily behaving").

We would like to consider the difference of "non-regulars" from other groups. They differ from others in low score in factor "regularity of daily rhythm", which consists of variables: "eating regularly", "sufficient hydration" and "sufficient amount of sleep"). This group does not try to live especially healthy. They have average score in factor "physical activity and healthy diet". But they do not live especially unhealthy. They are average in both factors of health-compromising behaviour. This group is average except the problem with paying attention to their basic needs. If we take into consideration the Baumeister et al. (2003) conclusion it resembles in some way the eating disorder. But the confirmation would require another study with personality inventory included.

Concerning the statistically significant but weak results of our study, there arise some problems about the method of self-esteem investigation. It regards self-deception and vulnerability of self-esteem tests, which are based on self-report. Paulhus (1986) says that in psychological assessment we aim at the most accurate description of some cognitive or behavioural attributes. Which is in case of self-reports endangered by the possibility of misrepresentation. We would be sceptical of self-reports of intelligence

and having a good personality is typically rated as more desirable than intelligence. Thus it seems dangerous to ignore the possibility that at least some respondents systematically misrepresent their own personality. On the basis of this statement we can never be sure, what we really measure (or how precisely) with only self-reported based instruments. This also includes SES and SLCS-R. The solution is to test the tendency to self-deception and impression management (e.g. using Paulhus's Balanced Inventory of Desirable Responding – BIDR).

Also Gerrard et al. (2000) deals with self-deception in relation of self-esteem and health risk behaviour. He comes to the conclusion, that individuals with high self-esteem who engage in risk behaviour often utilize a variety of cognitive strategies, that protect them from fully acknowledging their vulnerability to the potential negative consequences of their behaviour (e.g. they minimize their estimates of personal risk and overestimate the prevalence of the risk behaviour among their peers). It seems, that individuals with high self-esteem tend to self-deception because they need to maintain the good view of themselves, however individuals with low self-esteem have not such tendency because their self-esteem is already low.

Another problem lies in more complex reality of self-esteem. Baumeister et al. (2003) state that more instruments should be included into measuring self-esteem. High inner consistency of measure instruments of this construct can in fact conceal real heterogeneity of psychological processes belonging to high and low scores. He suggests that researchers interested in self-esteem should begin to pay more attention to narcissism (especially in Baumeister et al., 1996; Baumeister et al., 2000), self-deception, stability of self-esteem, and other distinctions within the broad category of self-esteem. Kernis (2003) proposes concept of "optimal self-esteem". His goal is to show, that optimal and high self-esteem are different from each other. High self-esteem can be fragile or secure depending upon the extent to which it is defensive or genuine, contingent or true, unstable or stable, and discrepant or congruent with implicit (nonconscious) feeling of self-worth.

Considering previous findings it is difficult to defend studying relationships of self-esteem with other constructs based simply on measuring self-esteem by a self-report, without any correction like self-deception and impression management detection and without considering other dimensions of the construct. By including additional instruments into measuring self-esteem we can probably obtain more accurate results in research of relationship between self-esteem and health-related behaviour.

References

Abernathy, T. J., Massad, L., & Romano-Dwyer, L. (1995). The Relationship between smoking and Self-Esteem. *Adolescence*, 120, 30, 899-907.

Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of Threatened Egoism to Violence and Aggression: The Dark Side of High Self-Esteem. *Psychological Review*, 103, 5-33.

Baumeister, R. F., Bushman, B. J., & Campbell, W. K. (2000). Self-Esteem, Narcissism, and Aggresion: Does Violence Result From Low Self-Esteem or From Threatened Egotism?. *Current Directions in Psychological Science*, 9, 26-29.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles?. *Psychological Science in the Public Interest*, 4, 1-44.

Blascovich, J., & Tomaka, J. (1991). Measures of self-esteem. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes*. New York: Academic. 115-160.

Blatný, M. (2001). *Sebepojetí v osobnostním kontextu*. Brno: Masarykova univerzita v Brně.

Blatný, M., Hřebíčková, M., Millová, K., Plháková, A., Říčan, P., Slezáčková, A., et al. (2010). *Psychologie osobnosti*. Praha: Grada.

Csémy, L., Lejčková, P., Sadílek, P., & Sovinová, H. (2006). *Evropská školní studie o alkoholu a jiných drogách (ESPAD), Výsledky průzkumu v České republice v roce 2003*. Praha: Úřad vlády České republiky.

DeHart, T., Tennen, H., Armeli, S., Todd, M., & Affleck, G. (2008). Drinking to regulate negative romantic relationship interactions: The moderating role of self-esteem. *Journal of Experimental Social Psychology*, 44, 527-538.

Gerrard, M., Gibbons, F. X., Reis-Bergan, M., & Russell, D. W. (2000). Self-Esteem, Self-Serving Cognitions, and Health Risk Behavior. *Journal of Personality*, 68, 1177-1201.

Glendinning, A., & Inglis, D. (1999). Smoking behavior in youth: the problem of low self-esteem?. *Journal of Adolescence*, 22, 673-682.

Hibell, B., Guttormsson, U., Ahlström, S., Balakireva, O., Bjarnason, T., Kokkevi, A., & Kraus, L. (2012). *The 2011 ESPAD Report*. Stockholm: modintryckoffset AB.

Kernis, M. H. (2003). Toward a Conceptualization of Optimal Self-Esteem. *Psychological Inquiry*, 14, 1-26.

Kounenou, K. (2010). Exploration of the relationship among drug use & alcohol drinking, entertainment activities and self-esteem in Greek University students. *Procedia Social and Behavioral Sciences*, 2, 1906-1910.

Paulhus, D. L. (1986). Self-Deception and Impression Management in Test Responses. In A. Angleitner, & J. S. Wiggins (Eds.), *Personality Assessment via Questionnaires.* Berlin, Heidelberg, New York, Tokyo: Springer-Verlag, 143-165.

Tafarodi, R. W., & Swann, Jr., W. B. (1995). Self-liking and self-competence as dimensions of global self-esteem: Initial validation of measures. *Journal of Personality Assessment*, 65, 322-342.

Ústav zdravotnických informací a statistiky. (2004). *Světové šetření o zdraví v České Republice [Elektronická verze]*. Praha: ÚZIS.

Ústav zdravotnických informací a statistiky. (2006). *Výběrové šetření o zdravotním stavu a životním stylu obyvatel České republiky zaměřené na zneužívání drog [Elektronická verze]*. Praha: ÚZIS.

The Motivations of Marathoners Scales -MOMS: Evaluation of motives for Greek swimmers with disabilities

Georgia Polatidou[1] , Sofia Batsiou [1], and Panagiota Polatidou[2]

[1]Democritus University in Thrace, Department of Physical Education and Sport Science
gpollati@phyed.duth.gr; smpatsio@phyed.duth.gr
[2]Aristotle University in Thessaloniki, Department of Physical Education and Sport Science
ppolatid@phed.auth.gr

Abstract. Sport psychologists have pointed out the importance of psychological skills training when it comes to disabled athletes. The study examined motives of Greek swimming disabled athletes with regard to their age, gender, classification, acquired or disability by birth, athletic experience and weekly training hours. Eighty two athletes (58 men 24 women), aged 10-59, (mean 31, s= 8) who participated in the National Swimming Championship, all classified S1- S10, answered voluntarily The Motivations of Marathoners Scales (MOMS) (Masters, Ogles, & Jolton, 1993). Analysis of reliability and factor analysis -principal components revealed the scale as a reliable and valid instrument that could be used to evaluate motives in Greek swimmers with physical disabilities. Non parametric analysis of variance Test Kruskal-Wallis and Test U Mann- Whitney were used and results of the research showed that disabled swimmers present highly developed ability of personal and social motivation. Female swimmers gave greater attention to internal psychological characteristics than male swimmers, while young swimmers up to 29 years of age developed a positive self-image and higher self-esteem compared with older swimmers. Classification influenced the motives of the meaning of life. The weekly training time affected the motives of target and social interaction. No significant differences were found on motivation when it came to the type of disability, acquired or disability by birth and athletic experience

Keywords: *motives, swimming, disabilities, classification*

Introduction

The need for greater research in sport psychology for disabled athletes has been documented within the bibliography (Page, Martin, & Wayda, 2001; Sherrill & Tripp, 2004). Sport psychologists have pointed out the importance of psychological skills training for athletes with physical disabilities (Hanrahan, 1998; Kasum, Lazarević, Jakovljević, & Bačanac, 2011; Martin, 2008; Sporner et al., 2009). Porretta and Moore (1997) proposed applied areas of psychological research that would be

beneficial for athletes with severe disabilities. Such areas are the identification of motivators, goal setting, self-regulation and visual rehearsal. Cox and Davis (1992) conducted a research, in which they compared psychological skills between elite wheelchair track or field athletes and able-bodied athletes. Other researchers examine the psychometric properties of athletic identity scale for swimmers with disabilities (Martin, Eklund, & Mushett, 1997). The characteristics of competitive anxiety of various athletes with disabilities have also been researched (Ferreira, Chatzisarantis, Gaspar, & Campos, 2007).

Sport psychologists use a variety of self-report inventories for the purpose of assessing an athlete's psychological characteristics and related behavior (Jagacinsk & Duda, 2001; Spencer, 2003). Based on the results of such psychological assessment, strategies or programs designed to remediate equivalent problems noted that psychological deficiencies should be implemented (Bawden, 2005). Valid and reliable material is essential both for the assessment of the athletes' psychological profile and for the development of psychological skills.

Some of the most valid and reliable evaluation scales of motivation which have been used in scientific studies are: Intrinsic Motivation Inventory (Ryan, 1982; McAuley, Dancan & Tammen 1989), Learning and Performance Orientations in Physical Education Classes Questionnaire (LAPORECQ) (Papaioannou, 1994), The Sport Motivation Scale (Pelletier et al.,1995), Exercise Regulations Questionnaire (BREQ-2) (Markland & Tobin, 2004), The Exercise Motivations Inventory - 2 (EMI-2) (Markland & Ingledew, 1997), Motives for Physical Activities Measure – Revised (MPAM-R) (Ryan, Frederick, Lepes, Rubio & Sheldon, 1997), Motivations of Marathoners Scales (MOMS) (Masters, Ogles & Jolton, 1993).

There has been an increase in the amount of research on sport psychology in Greece (Georgiadis, 2009; Diamantakos, 2008) however, studies involving psychological descriptions of competitive and noncompetitive athletes with physical disabilities have received much less attention. Therefore, in the past twelve years, Greek researchers have been focusing on research both at national (Maggouitsa et al., 2004) and international level athletes with physical disabilities (Kokaridas, Natsis, Markopoulos, Chatzigeorgiadis, & Karpathakis, 2005; Skordilis, Koutsouki, Asonitou, Evans, & Jensen, 2002) in order to help improve their exercise programs based on their psychological characteristics such as sport orientations, athletic identity, sport motivation and goal perspectives. In addition, Goudas, Kontou, and Theodorakis (2006) have developed a Greek version of the Test of Performance Strategies (TOPS), indicating with their results that this version could be used for assessing psychological skills for the disabled Greek athletes. Another valid and reliable tool for the Greek population of disabled athletes, the motives scale MOM's, has revealed a useful asset for coaches and sport psychologists (Polatidou, Batsiou, & Polatidou, 2010). MOM′S scale was used instead of another motive scale initially due to proposal of the creators of MOM′S scale for adaptation in special populations and thereafter to some common character traits of disabled athletes compared with marathon runner athletes. These characteristics have to do with the centrality of what they do in their lives. They must train more than is needed, expose themselves to stress and strain well beyond what is necessary for any other sport athletes, firstly to maintain fitness and to overcome

disability boundaries, apart from the obvious physical and psychological effort, cancel or postpone engagements, alter work and eating schedules and continue regimens for extended periods that may provide a model of "super- adherence" (Dishman, 1988). Therefore, Greek researchers pointed out the importance of evaluating psychological skills of entire population of Greek disabled competitive athletes in all sports, particularly in swimming were with Athletics are the most popular sports in Greece.

The purpose of this study was to evaluate a) the internal consistency and structural validity of The Motivations of Marathoners Scales –MOMS (Masters, Ogles, & Jolton, 1993) for the Greek population of disabled swimmers b) the motives of Greek competitive swimmers with physical disabilities and c) to examine the differences of motivation among Greek competitive swimmers with disabilities with regard to gender, age, classification, acquired or disability by birth, athletic experience and weekly training hours.

Data and Methodology

Participants

All swimmers were informed about the purpose of the study and completed voluntarily and anonymously the questionnaire. The study was part of a doctoral thesis which was approved by the Department of Psychical Education and Sports Sciences ensuring the privacy of personal information of athletes and guaranteeing institutional ethical approval. Eighty two (n=58 men, n=24 women) disabled competitive swimmers (from the total amount of 121 participants of the National Championship) participated voluntarily in the study. Their Classification Status ranged S1- S10: (nS1=1, nS2=6, S3=7, nS4=2, nS5=7, nS6=16, nS7=13, nS8=11, nS9=8 and nS10=10). Thirty nine of the participants had acquired physical disability and 43 of them disability by birth. Swimmers had various disabilities: Cerebral Palsy (n=32), Spinal Cord Injury (n=18), Amputations (n=8), Poliomyelitis (n=4), Achondroplasia (n=3), Traumatic Brain Injury (n=2), Multiple Sclerosis (n=2), Muscle Dystrophy (n=1), orthopedic disorders and rheumatic diseases (n=12). Researchers divide them in five age groups: 10-19 years (n=13), 20-29 (n=22), 30-39 (n=36), 40-49 (n=8) and 50-59 (n=3). All swimmers were also divided by athletic experience: 1-3 years (n=11), 4-6 (n=27), 7-10 (n=15), over 11 years (n=29) and by weekly training hours: six hours per week (n=24), eight (n=18) and ten (n=40).

Measures

The MOMS, which was adapted to the population of Greece by Polatidou, Batsiou, and Polatidou, (2010) was used to measure motivation in sport of swimming for disabled athletes. This particular measure of coping is composed by the following general categories and each subscales: 1. Physical Health Motives: General Health Orientation, Weight Concern, 2. Social Motives: Affiliation, Recognition, 3. Achievement Motives: Competition, Personal Goal Achievement, 4. Psychological Motives: Psychological Coping, Self-Esteem, Life Meaning. All items of the subscales of the MOMS were assessed on a 7-point Likert scale ranging from 1 (*not a reason*) to 7 (*very important reason*).

Procedure

The MOMS was given to the swimmers at a regular swimming training in the middle of training season. Instructions were read by the investigator or by the coach and athletes were informed that participation was voluntary and that all the information reported in their questionnaire was confidential. Fifty two swimmers from 11 clubs in Northern Greece completed the questionnaire under the guidance of the researcher inside swimming facilities of the town each athletic club cited. Thirty swimmers from 11 clubs in Southern Greece completed the questionnaire after hearing the written instructions from their coaches at the swimming facilities of the town each athletic club cited and these questionnaires returned to the researcher by post. All participants signed and returned an informed consent form and were instructed to consider each item and answer without consulting any other individuals. Instructions were the same for all participants. In the last page of the questionnaire participants were instructed to fill out a brief demographic summary: age, gender, classification, acquired or disability by birth, type of disability, athletic experience and weekly training hours.

Statistical analyses

The Statistical Package for the Social Sciences (SPSS) version 18.0 was used as a tool for data analysis. Statistical analyses included descriptive statistics for demographic data, reliability and factor analysis of the MOMS, non-parametric analysis of variance Test Kruskal-Wallis for the variables of age, classification, type of disability, years of athletic experience and weekly training hours and Test U Mann- Whitney (2 independent) for the variables of gender and acquired or disability by birth.

Empirical Analysis

Table 1. presents the descriptive statistics for factor composite scores and Cronbach's' alpha for the nine Subscales of the questionnaire. From the main scores it appears that eight out of nine subscales tend to reach the maximum score (7 –very important reason) for each question (range 5,29 to 6,10). Respectively one subscale exceed the half of maximum score (3,92).   Cronbach's Alpha indicator proved to be satisfying to high as far as the reliability of the scale (.920) and of the 9 subscales (.919-.987) are concerned in the questionnaire responses showed that the Cronbach' alpha was extremely high for both total scale and nine subscales (Howitt & Cramer, 2006).

Table 1.: Means, Standard Deviation and Cronbach's' alpha for MOMS and Subscales

Subscales	N	M ±SD	Cronbach's alpha
Weight Concern	4	3,92±.1,38	.934
Competition	4	5,29± 1,05	.946
Recognition	6	5,87±..81	.942
Personal Goal Achievement	6	5,97±.70	.921
Affiliation	6	6,10±.81	.958

General Health Orientation	6	5,84±.77	.924
Psychological Coping	9	5,33±1,19	.987
Self-Esteem	8	5,80±.86	.984
Life Meaning	7	5,36 ±.80	.919
Total scale	**56**		**.920**

Factor analysis (main axes- principal components) of the replies of athlete's revealed nine factors with an eigenvalue over 1 explaining 84,78% of the total variance. Test of Kaiser-Meyer-Olkin as a measure of sampling adequacy (KMO) and Bartett's sphericity Test were conducted revealing high values (range .77 to .93) for all nine factors (Table 2). All items of each factor presented a range of medium to high loads (Table 3).

Table 2.: KMO Test and the test of Bartett's sphericity

Subscales	KMO	Bartett's Test	
		Chi-Square	P
Weight Concern	.84	279,486	**,000**
Competition	,84	311,527	**,000**
Recognition	,87	597,281	**,000**
Personal Goal Achievement	,83	382,037	**,000**
Affiliation	,83	661,037	**,000**
General Health Orientation	,83	487,645	**,000**
Psychological Coping	,93	1721,069	**,000**
Self-Esteem	,90	1358,103	**,000**
Life Meaning	,77	642,965	**,000**
Total Scale	,59	7197,633	**,000**

Table 3.: Components in Rotation axes (varimax).

					Components				
Subscales									
	1	2	3	4	5	6	7	8	9
Psych. Coping 9*	,959								
Psych. Coping 8	,958								
Psych. Coping 2	,955								
Psych. Coping 3	,951								
Psych. Coping 1	,951								
Psych. Coping 4	,949								
Psych. Coping 5	,947								
Psych. Coping 6	,929								

Psych. Coping 7	,795					
Self-Esteem 3		,968				
Self-Esteem 6		,967				
Self-Esteem 2		,966				
Self-Esteem 1		,952				
Self-Esteem 5		,947				
Self-Esteem 4		,917				
Self-Esteem 8		,850				
Self-Esteem 7		,840				
Affiliation 4			,887			
Affiliation 1			,869			
Affiliation 5			,869			
Affiliation 2			,828			
Affiliation 6			,739			
Affiliation 3			,706			
Recognition 2				,938		
Recognition 5				,933		
Recognition 4				,915		
Recognition 3				,834		
Recognition 6				,771		
Recognition 1				,655		
Life Meaning 4					,901	
Life Meaning 5					,890	
Life Meaning 6					,785	
Life Meaning 2					,749	
Life Meaning 1					,738	
Life Meaning 3					,724	
Life Meaning 7					,719	
Health 4						,899
Health 6						,897
Health 5						,894
Health 3						,797
Health 2						,795
Health 1						,721
Goal 4						,857
Goal 5						,822
Goal 6						,769
Goal 1						,763
Goal 3						,745

Note: "Health" and "Goal" values appear in the rightmost columns; Health 4–1 occupy the sixth factor column and Goal 4–3 occupy the seventh factor column as shown in the image.

Goal 2	,744	
Competition 2	,895	
Competition 4	,861	
Competition 3	,851	
Competition 1	,842	
Weight Concern 3		,908
Weight Concern 2		,875
Weight Concern 4		,826
Weight Concern 1		,810

*Note: *The number next to the subscales indicates the order of appearance the issue of each factor on the scale.*

Non parametric test U of Mann-Whitney (2 independent) for variable "gender" found significant difference in subscale "Psychological Coping" (U=500,500 N1=58, N2=24, z=--1,997, *P* <.05) (Table 4). By checking the Mean Rank women scored significantly higher than men (Table 5). Non parametric test U of Mann-Whitney revealed no statistically significant difference between the averages of the factors of MOMS and "acquired or disability by birth" (Table 4).

Table 4.: Test U Mann- Whitney for gender and acquired or disability by birth

	Gender			**acquired/ disability by birth**		
	U	**Z**	**P**	**U**	**Z**	**P**
Weight Concern	602,500	-,955	,339	793,500	-,419	,675
Competition	571,000	-1,281	,200	778,500	-,560	,575
Recognition	681,500	-,149	,882	677,000	-1,508	,131
Goal	658,000	-,390	,697	765,000	-,687	,492
Affiliation	620,500	-,749	,454	755,000	-,781	,435
Health	620,500	-,774	,439	690,500	-1,382	,167
Psychological Coping	500,500	-1,997	**,046***	799,000	-,368	,713
Self-Esteem	626,500	-,730	,466	748,000	-,866	,387
Life Meaning	546,000	-1,532	,125	799,000	-,368	,713

*Note: *statistical significance level P<.05*

Non parametric test Kruskal-Wallis for the variable "age" also found significant differences in the subscale "Self-Esteem" (x^2=14,549, df=4, *P* = .006). In order to find separate differences, the U test of Mann Whitney was conducted after applying Bonferroni correction, which resulted in coefficient significance *P* =0,005. By checking the Mean Rank, the results revealed that athletes 10-19 and 20-29 years old

scored significantly higher on the parameter Self-Esteem than athletes 30-39 years old (Mean Rank) (Table 5).

Non parametric test Kruskal-Wallis for the variable "weekly training hours" also found significant differences in the subscale "goal" (x^2=6,588, df=2, P =,037), "Affiliation" (x^2=9,784, df=2, P =,008), "Psychological Coping" (x^2=7,667, df=2, P =,022) (Table 6). In order to find separate differences, the U test of Mann Whitney was conducted after applying Bonferroni correction, which resulted in coefficient significance $P = 0,017$. By checking the Mean Rank, the results revealed that athletes with six hours of training per week scored significantly higher on the parameter of goal and affiliation than those with ten hours of training and athletes with eight hours of training per week scored significantly higher on the parameter Psychological Coping than those with six and ten hours of training (Mean Rank) (Table 5). There was no statistically significant difference between the averages of the factors of MOMS and the variables of "type of disability", "years of athletic experience" (Table 6).

The comparison between ten categories of classification status (S1-S10) revealed significant difference in the subscale "Life Meaning", (x^2=17,955, df=9, P =.036). In order to find separate differences, the test U of Mann Whitney was conducted after applying Bonferroni correction, which resulted in coefficient significance $P = 0,001$. Analysis in this level of coefficient significance reveals no statistical differences between pairs of comparisons of classes. Values of Mean Ranks for each Class are available in Table 7.

Table 5.: Mean Ranks in gender, age and h/week comparisons

Subscales	Gender	N	Mean Rank	Subscale	age	N	Mean Rank	Subscales	h/week	N	Mean Rank
Psychological Coping	M	58	38,13	**Self-Esteem**	10-19	25	30,42	**Goal**	6	24	39,79
	F	24	49,65		30-39	23	18,07	**Affiliation**	10	40	28,13
					20-29	23	28,98		6	24	40,63
					30-39	23	18,02	**Psychological Coping**	10	40	27,63
									6	24	17,42
									8	18	26,94
									8	18	37,75
									10	40	25,79

Table 6.: Kruskal-Wallis test for age, h/week, classification, type of disability and years of athletic experience

	Age			h/week			Classification			type of disability			years of athletic experience		
	X^2	Df	P	X^2	Df	P	X^2	Df	P	X^2	Df	P	X^2	Df	P
Weight Concern	1,138	4	,888	1,646	2	,439	3,832	9	,922	4,088	8	,849	1,797	3	,616
Competition	1,465	4	,833	1,850	2	,396	8,027	9	,531	4,749	8	,784	3,285	3	,350
Recognition	6,953	4	,138	,783	2	,676	10,442	9	,316	6,864	8	,551	5,636	3	,131
Goal	3,041	4	,551	6,588	2	**,037***	14,251	9	,114	4,717	8	,787	4,405	3	,221
Affiliation	4,667	4	,323	9,784	2	**,008***	4,061	9	,907	1,826	8	,986	1,982	3	,576
Health	,741	4	,946	,102	2	,950	13,558	9	,139	3,574	8	,893	1,046	3	,790
Psychological Coping	1,181	4	,881	7,667	2	**,022***	9,256	9	,414	5,948	8	,653	5,812	3	,121
Self-esteem	14,549	4	**,006***	3,041	2	,219	7,474	9	,588	5,513	8	,702	6,455	3	,091
Life meaning	2,876	4	,579	4,917	2	,086	17,955	9	**,036***	6,248	8	,619	1,910	3	,591

Note: *statistical significance level $P<,05$

Table 7.: Mean Ranks for classification status comparisons

Subscale	classification	N	Mean Rank
Life meaning	1	2	68,50
	2	6	26,75
	3	7	35,57
	4	2	69,50
	5	7	47,36
	6	16	42,38
	7	13	45,77
	8	11	24,73
	9	8	56,88
	10	10	38,60
	Total	**82**	

Conclusions/Discussion

Reliability analysis (Cronbach's a) showed that there was strong internal consistency among the questions regarding the nine subscales. Factor analysis confirmed the structural validity of the scale, proving that it can be a reliable tool concerning its use on Greek swimmers with physical disabilities. Respectively, high levels of internal consistency and construct validity were observed in all subscales of studies that used the scale in able-bodied athletes (Havenar & Lochbaum, 2007; Krouse, Ransdell, Lucas & Pritchard, 2011). Therefore, the suggestions for adapting the scale in special population from the researchers who constructed the scale are confirmed (Masters, Ogles & Jolton, 1993).

   High scores in most of questions of the questionnaire factors revealed that swimmers with physical disabilities were highly motivated from social and personal achievement motives. According to the Theory of Iinstrumentality, one of the bases motivation theories of research psychology, the main force behind an act is affected by the consequences of that act (Heckhausen, 1989). In the present study personal and social motives of disabled swimmers developed (positively affected) as they exercise in a competitive level and through their participation in swimming meetings. Additionally, the scores of swimmers with disabilities are significantly higher than the corresponding able-bodied athletes (Havenar & Lochbaum, 2007) except from the weight concern motive in which able- bodied athletes scores higher (Choate, 2008). This difference is explained from the fact that disabled athletes through their engagement with sports understand more effectively the physiological processes caused by exercise, that results in regaining the lost sense of their body and take full control of it, without focusing on body weight (Costa, 1992). This physical control improves their self-image and self-esteem in a way that they have enough independence to strive their reintegration into society and the activities of their daily lives. It leads naturally to learn their limits and optimize their abilities to deal with the

all problems, psychologically or physically and finally overcome them. Also, the psychological treatment in situations of anxiety, depression and anxiety occurs more comprehensive compared with able-bodied athletes (Perreault & Marisi, 1997).

Results concerning gender indicate that women had significantly higher scores on the factor of Psychological Coping. This fact is confirmed by studies that continually highlight the importance of the experience of participation in the sport of women with disabilities in both social and personal development and pointed out that women are guided by intrinsic motivation towards sport in contrast to men, who are motivated by extrinsic reasons (Papanastasiou, 2009). As far as age is concern it turns out that young disabled swimmers between 10- 29 of years of age developed a high positive self-image and higher self-esteem compared with older athletes, through the process of dealing with the sport of swimming. Respectively, research that supports organizing and structuring exercise programs for young people with disabilities, showed better level of self-concept positive self-image and higher self-esteem of young people with disabilities compared with older disabled people (Ioannidou, Batsiou, Douda, & Kourtessis, 2011).

The results of weekly training hours display special interest. The diversity of every athlete in combination with the demanding environment of swimming training for athletes with disabilities (access to swimming facilities, entrance and exit the pool, locker room-wc) may affect the parameter of "goal achievement" and "affiliation" which leads swimmers who train fewer hours - not by choice but by external factors – to achieve higher scores than those with more hours of training. Lunkes & Faria (2007) have pointed out the importance of affiliation motives that presented the highest scores while evaluating motives in wheelchair basketball players. The dominant feature was the "new friends" and "the way out of the home environment". They also pointed the importance of the disabled athlete's autonomy as a mean in managing psychological and social independence, as most of them live with their families and depend upon them, thus follow a restrictive lifestyle.

Respectively, in the present study the functional ability of swimmers determined by the classification status affected the score of swimmers in the parameter "life meaning" without though pointing any difference between classes. This was may be related to their already developed interest in life and a better quality of life of athletes with disabilities who are already involved in physical activity and a sport, regardless of the degree of functionality, motor limitations and severity of disability (Tasiemski, Kennedy, Gardner & Blaikley, 2004).

No significant differences were found on the average of subscales compared with acquired or disability by birth, type of disability and athletic experience. The aim of sports for disable is the interaction of people with different disabilities and able-bodied, the increase of successful experiences, the improvement of confidence and general condition of individuals, the prevention of possible relapse of an existing disability. So, motives in this study are reviewed and evaluated under the general spectrum of disability without presenting differences according to different types of disabilities.

Regarding acquired or disability by birth studies point out the importance of exercise and sports involvement both in psychological motives and in lifelong involvement with physical activity (Furst, Ferr & Megginson, 1993; Gürsel & Koruç, 2011). Finally, this study did not reveal differences between more or less experienced

swimmers thus giving an indication that motivation of swimmers with physical disabilities is a lifelong sports feature, which is not associated with their racing experiences. In this way, according to Furst, Ferr & Megginson (1993), the unique dimension of motives for continuous training and not the occasional involvement with the sport is highlighted**.**

   Further studies are necessary in order to adapt other psychological scales to the population of Greek disabled athletes, to evaluate psychological characteristics of Greek disabled athletes in all Paralympic Sports and to also compare the differences between sports.

References

Bawden, M. (2005). Providing Sport Psychology Support for Athletes with Disabilities. In Dosil, J. (Ed.), *The Sport Psychologist's Handbook: A Guide for Sport-Specific Performance Enhancement* (pp. 665-683). Chichester, West Sussex, UK: John Wiley & Sons Ltd.

Choate, M. (2008). Marathon Participation: A Study of Motivation and Training Habits of Runners Affiliated With a Charitable Cause. *United States Sports Academy.* http://books.google.gr/books?id=xDCzOZVdEPQC&printsec=frontcover#v=onepage &q&f=false

Cox, R., & Davis, R. (1992). Psychological skills of elite wheelchair athletes. *Palaestra, 8(3),* 16-28.

Diamantakos, I. (2008). Special psychological aspects of training Basketball players. Retrieved from http://www.sepk.gr/portal2/el-GR/-basketball-coach-topics/sports-psychology/4043-2011-05-13-06-03-22

Dishman, (1988). In Masters, K. S., Ogles, B. M., & Jolton, J. A. (1993). The development of an instrument to measure motivation for marathon running: The motivations of marathoners scales (MOMS). *Research Quarterly in Exercise and Sport, 64*, 134-143

Georgiadis, E. (2009). Psychological preparation and athletic performance. Retrieved from http://www.psychologyonline.gr/default.aspx?id=52&parent=50.

Goudas, M., Kontou, M., & Theodorakis, Y. (2006). Validity and reliability of the Greek version of the Test of Performance Strategies (TOPS) for athletes with disabilities. *Japanese Journal of Adapted Sport Science*, *4,* 29-36.

Gürsel, F., & Koruç, Z. (2011). The influence of physical activity on body image in people with and without acquired mobility disability *Acta Universitatis Palackianae Olomucensis. Gymnica, 41, 29-35.*

Hanrahan, S. J. (1998). Practical considerations for working with athletes with disabilities. *The Sport Psychologist, 12*, 346 357.

Havenar, J., & Lochbaum, M. (2007). Differences in Participation Motives of First-time Marathon Finishers and Pre-Race Dropouts. *Journal of Sport Behavior. 30,* 270-279.

Heckhausen, H. (1989). In Μπάουμαν, Σ., & Τσορμπατζούδης, Χ. (2007). Ψυχολογία στον αθλητισμό [Sports Psychology]. Thessaloniki: Copy City.

Ioannidou, D., Batsiou, S., Douda, H. & Kourtessis, T. (2011). The Self- Perception level of children with mobile disabilitied and what motivates them to exercise. Short papers presented during the 19th International Congress of Physical Education and Sport. http://www.phyed.duth.gr/files/congress/2011/spaper/Exerc_disabilities.pdf

Jagacinski, C., M. & Duda, J., L. (2001). A comparative analysis of contemporary achievement goal orientation measures. *Educational and Psychological Measurement, 61*, 1013- 1039.

Ferreira, J. P., Chatzisarantis, N., Gaspar, P. M., & Campos, M. J. (2007). Precompetitive anxiety and self confidence in athletes with disability. *Perceptual and Motor Skills*, *105*, 339-346.

Furst, D., Ferr, T., & Megginson, N. (1993). Motivation of disable athletes to participate in triathlons. *Psychological Reports, 72,* 403-406.

Kasum, G., Lazarević, L., Jakovljević, S., & Bačanac, L. (2011). Personality of male Wheelchair Basketball players and nonathletes individuals with disability. *Physical Education and Sport 9*, 407 – 415.

Kokaridas, D., Natsis, P., Markopoulos, K., Chatzigeorgiadis, A., & Karpathakis, N. (2005). Sports orientation and athletic identity of Shooters in Paralympic Games in Athens. *Inquiries in Physical Education and Sport*, *3*(1), 98-106. Retrieved from http://www.hape.gr

Krouse, R., Ransdell, L.B., Lucas, S., & Pritchard, M. (2011). Motivation, goal orientation, coaching and training habits of women ultrarunners. *Journal of Strength and Conditioning Research, 25,* 2835-2842.

Maggouritsa, G., Kokaridas, D., Chassandra, M., Perkos, S., Tsapatos, S., & Davos., A. (2004). Sports orientation in Basketball wheelchair athletes. *Inquiries in Physical Education and Sport*, *2*(1), 67-72. Retrieved from http://www.hape.gr

McAuley, E., Dancan, T. E., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A confirmatory Factor Analysis. *Research Quarterly for Exercise and Sport, 60,* 48-58.

Masters, K. S., Ogles, B. M., & Jolton, J. A. (1993). The development of an instrument to measure motivation for marathon running: The motivations of marathoners scales (MOMS). *Research Quarterly in Exercise and Sport, 64*, 134-143.

Markland, D., & Ingledew, D. K. (1997). The measurement of exercise motives: Factorial validity and invariance across gender of a revised Exercise Motivations Inventory. *British Journal of Health Psychology, 2,* 361-376.

Markland, D., & Tobin, V. (2004). A modification of the Behavioral Regulation in Exercise Questionnaire to include an assessment of motivation. *Journal of Sport and Exercise Psychology, 26,* 191-196.

Martin, J.J. (2008). Multidimensional Self-Efficacy and Affect in Wheelchair Basketball Players. *Adapted Physical Activity Quarterly, 25*, 275-88.

Martin, J. J., Eklund, R. C., & Mushett, C. A. (1997). Factor structure of the Athletic Identity Measurement Scale with athletes with disabilities. *Adapted Physical Activity Quarterly*, *14,* 74-82.

Lunkes, L., & Faria, C. C. (2007). Motives that influenced the start of the practice of wheelchair basketball. *Sobama Journal, 12 (1):* 140-143.

Page, S. J., Martin, S. B., & Wayda, V. K. (2001). Attitudes toward seeking sport psychology consultation among wheelchair basketball athletes. *Adapted Physical Activity Quarterly*, *18,* 183–192.

Papaioannou, A. (1994). Development of a questionnaire to measure achievement orientations in physical education. *Research Quarterly for Exercise and Sport, 65 (1),* 11-20.

Papanastasiou, P. (2009). Ο ελεύθερος χρόνος και τα κίνητρα, ως όπλο της διαφήμισης και προβολής στα κέντρα άσκησης και αναψυχής. [*Leisure time and*

motivation, as tool of advertising and promotion in fitness and leisure centers]. Undergraduate Thesis*, F*aculty *of Management and Economics* of Crete http://nefeli.lib.teicrete.gr/browse/sdo/mk/2009/PapanastasiouPolikseni/attached-document-1307013201-251679-14020/papanastasiouPolikseni.pdf

Pelletier, L. G., Fortier, M. S., Vallerand, R. J., Tuson, K. M., Brière, N. M., & Blais, M. R. (1995). Toward a new measure of intrinsic motivation, extrinsic motivation, and amotivation in sports: The Sport Motivation Scale (SMS)*. Journal of Sport & Exercise Psychology, 17,* 35-53.
Perreault, S., & Marisi, D. (1997). A test of Multidimensional Anxiety Theory with male wheelchair basketball players. *Adapted Physical Activity Quarterly, 14,* 108-118.

Polatidou, G., Batsiou, S., & Polatidou, P. (2010, December). Εγκυρότητα και αξιοπιστία της κλίμακας MOMS για τους Έλληνες αθλητές με κινητικές αναπηρίες [*Validity and reliability of The Motivations of Marathoners Scale- MOMS for Greek athletes with disabilities*]. Full Paper presented at the Eleventh National Theme Conference of Northern Greece Physical Education Teachers Association, Thessaloniki.

Porretta, D. L., & Moore, W. (1997). A review of sport psychology research for individuals with disabilities: Implications for the future. *Clinical Kinesiology*, *50,* 83-93.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43,* 450-461.

Ryan, R. M., Frederick, C. M., Lepes, D., Rubio, N., & Sheldon, K. M. (1997). Intrinsic motivation and exercise adherence. *International Journal of Sport Psychology, 28,* 335-354.

Skordilis, E., K., Koutsouki, D., Asonitou K., Evans E., & Jensen, B. (2002). Comparison of sport achievement orientation between wheelchair and able-bodied Basketball athletes, *Perceptual and Motor Skills, 94*(1), 214-218.

Sherrill, C., & Tripp, A., (2004). Self-concept, motivation and well-being. In Sherrill, C. (Eds.), *Adapted Physical Activity: Recreation, and Sport* (p.p.218-239) New York: McGraw Hill.

Sporner, M. L., Fitzgerald, S. G., Dicianno, B. E., Collins, D., Teodorski, E., Pasquina, P. F., & Cooper, R. A. (2009). Psychosocial impact of participation in the National Veterans Wheelchair Games and Winter Sports Clinic. *Disability & Rehabilitation, 31*, 410-18.

Tasiemski, T., Kennedy, P., Gardner, B.P., & Blaikley, R. A. (2004). Athletic identity and sports participation in people with spinal cord injury. *Adapted Physical Activity Quarterly, 21,* 364-378.

Parental Influence on Children's Nutrition and Physical Activity Behaviours in Greece

George Pavlidis[1] and Angelos Rodafinos[2]

[1] University of Sheffield (SEERC), IF,
gipavlidis@seerc.org
[2] Swinburne Online Education, Melbourne
arodafinos@swinburneonline.com

Abstract. Obesity places psychological, social, and physical challenges to children and constitutes a serious public and financial concern. Previous research findings highlighted the influence of parents' nutrition and physical activity, as well as their knowledge on these domains, on those of their children. This study was set out to examine if parental weight, nutrition and physical activity behaviours, as well as their knowledge in these domains, are related to their children's. A total of 306 children and parents completed a survey assessing weight, diet, PA and knowledge in these domains. Results indicated parental weight, nutrition behaviours, and nutrition knowledge, to be positively correlated with those of their children. Physical activity behaviours were correlated only between daughters and their fathers. Parental breakfast consumption and mothers weight status was predictive of children's weight status. It appears that parental nutrition behaviour and knowledge may influence those of their children.

Keywords: *nutrition, physical activity, parents, children, obesity*

1   Introduction

The contemporary view of health is not limited in the notion of the mere absence of physical illness, but is defined as a positive state of physical, mental, and social well-being (Sarafino, 2008). Following this notion for health, children's health status in western contemporary societies is in jeopardy. This is colourfully depicted by recent reports on children's weight in western societies, which indicate that 19.8% of children between the ages of 6-19 in the United States are obese (Center of Disease prevention and Control [CDC], 2011). In Europe 30% of children is overweight, from which 10% fulfil the criteria of obesity (Lobstein, 2003). The figures are similarly worrying in Greece, where 26% of boys and 19% of girls are overweight, from which 11% fulfil the criteria of obesity (Chaniotis, Botsari, Mikelopoulou, & Chaniotis, 2010). The obesity rates among youths are increasing; the relevant rates doubled in

10 years for the 12-19 age groups and tripled for the 6-11 age group (Zapata, Bryant, Macdermott, & Hefelfinger, 2008).

It is widely observed and accepted that obese children face certain psychological, social, and physical challenges, such as physical appearance anxiety (Kring, Johnson, Davison, & Neale, 2010); low self-esteem, negative self-image, and depressive symptoms (Koplan, Liverman, & Kraak, 2005); bullying (Gray, Kahhan & Janicke, 2009; Strauss & Pollack, 2003; Janssen, Craig, Boyce, & Pickett, 2004). Jain (2004) argued that 40 to 70% of obese children evolve to be obese adults. In turn, obese adults face higher rates of premature death (Doak, Vissher, Renders, & Seidell, 2006). Obesity has been also related with rapid increase of obesity-specific hospital expenditures (BBC, 2006; Koplan et al., 2005; Wang & Dietz, 2002). These expenditures include a) the direct health-care costs of chronic diseases related to obesity and b) the indirect "external costs" that result from the sedentary life of obese individuals. To put these costs in perspective, Sturm (2002) posed that the health-care costs of obesity are greater than those of the smoking-related and alcohol consumption-related diseases.

This obesity epidemic motivated the research community to explore the underlying mechanisms of surplus weight, as well as the relation between weight status, caloric intake, and energy expenditure. Caloric intake refers to the amount and the quality of food that people consume. Energy expenditure includes the amount of energy that vital organs spend in order to function, the energy spent to metabolize food, and energy spent in any form of physical activity (PA) (Cutler, Glaeser, & Shapiro, 2003). A common conclusion was that in order to maintain a healthy weight status, the caloric intake has to be in equilibrium with one's amount of energy expenditure (Birch & Fisher, 1998; Chaniotis et al., 2010; Cutler et al., 2003).

Early findings indicated that obesity has a genetic basis that shapes a predisposition to eat high caloric food (Birch & Fisher, 1998; Benton, 2004; Breen, Plomin, & Wardle, 2006). Genes however, could not explain the obesity epidemic completely; obesity rates multiplied the last three decades, whereas natural selection cannot operate that quickly to shape genetic sequences (Welk, Wood, & Morss, 2003). Chu, Grossman, and Saffer (2004) posed an alternative explanation. They argued that the rapid economical and societal changes in the modern western society may have given easy access to pleasantly flavoured high-energy food and simultaneously may have reduced the demands in calories expenditure. If an inherited preference toward high caloric food is a fact, than the contemporary social context may predispose further to a malicious balance between caloric intake and energy expenditure in contemporary societies.

Accordingly, there is a broad consensus that children's PA and eating habits need to be approached from a social and ecological perspective. Children operate in the social environment and their behaviour is subjected to social influences derived from the media, peers, and parents (Benton, 2004; Birch & Fisher, 1998; Kral & Faith, 2009). Pratt (1976) explored the idea of a well-functioning family system that influences the health behaviours of children. Skinner, Carruth, Bounds, and Ziegler

(2002) stated that mothers influence their children's food preferences through their own preferences, which may also border the available food offered in the household (Koui & Jago, 2008). Similarly, evidence suggests that students' nutrition style is related to the control their parents exert on the quality and quantity of food in the house (Risvas, Panagiotakos, & Zampelas, 2008).

However, parents and children do not spend the whole day together, and to date most children in the western societies can afford to buy "prohibited" groceries out of their weekly allowance in their unsupervised time (Welk et al., 2003). Yet, family may remain a key system that impacts behaviour-specific beliefs and enforces children's health behaviours (Ornelas, Perreira, & Ayla, 2007). Horn and Horn (2007) stated that parental values and beliefs determine their behaviour towards their children, which in turn affects children's development, belief system, and PA behaviour. This may be through observational learning, namely the learning mechanism where children tend to imitate parental behaviours. As Bandura (1977) colourfully illustrated in the social learning theory, children's behaviour can be shaped through role modelling, social influence or social support. Notably, the majority of kids report that the person they would like to be the most is their father or their mother; daughters reported their mother and sons their father (American Dietetic Association Foundation, 2011). Parental role modelling therefore, could influence children's attraction and engagement to favourable/unfavourable amounts of PA and good/maladaptive nutrition habits (Risvas, Panagiotakos, & Zampelas, 2008; Welk et al., 2003).

There is empirical evidence indicating family members' physical activity or inactivity to be related (Fogelholm, Nuuttinen, Pasanen, Myohanen, & Saatela, 1999). Bois et al. (2004) showed both parental beliefs and their PA levels to be significantly related to those of their children's. Brown and Odgen (2004) found that parents and primary school students' consumption of healthy snacks is positive correlated. Malinadretos et al. (2009) and Kosti et al. (2008) posed that Greek children's weight is related to their parents'. Tsamita and Karteroliotis (2008) indicated that the consumption of unhealthy food of students relates to this of their parents. Similarly, Lazarou, Kalavana, and Matalas (2008) illustrated that parents' poor quality of nutrition is related to these of their children. Razakou, Tsapakidou, Besi, & Tsompanaki (2003) concluded that obese children are more probable to have parents who do not engage in any form of PA.

The American Dietetic Association Foundation (ADAF, 2011) advocated that parents should act as good exemplars in nutrition and PA. The Golan and Crow (2004) study depicted childhood obesity interventions that target only parents as agents of children's dietary and PA change to be superior from conventional approaches (such as targeting directly to children's behavioural alteration). Such interventions include parents' education on a) the importance of food availability in the household and the influence that it exerts to children's food preference (in quality and quantity) and b) the importance of parental role modelling in nutrition and PA behaviours (Golan & Crow, 2004).

It is evident that the examination of nutrition and PA practices within the family system is important in order to analyse the "systemic" influences that shape children's diet, PA patterns, and weight. Yet, to our knowledge there are no studies that investigate the relations between nutrition, PA, and weight between family members, in order to identify significant predictors of children's weight status. Therefore, this study was set out to examine whether a) parental diet and PA behaviours b) parental knowledge of optimal nutrition and PA c) parental weight status, relate to their children's diet, PA, knowledge of optimal nutrition and PA, as well as weight status. It was hypothesized that parental behaviour and knowledge regarding optimal nutrition and PA will be positively correlated with those of their children. Furthermore, it was hypothesized that parental weight status, diet and PA behaviours, relate to those of their children's.

2   Method

Participants

A total of 153 children (M_{age} = 11.88, SD = 2.24, 64% females) from a northern Greek City and one of their parents (M_{age} = 41.49, SD = 5.65, 72% females) participated in this study. Most of the participants were native Greeks. The daughters-mothers pairs that completed the questionnaires were 75; the sons-fathers pairs 20; the daughters-fathers 22; and the sons-mothers 34.

Instruments

Two open-ended self-report questionnaires were used in this study, one for children and one for parents.

Children's questionnaire.

The items that composed the questionnaire for children were taken directly from the Youth Physical Activity and Nutrition Survey (YPANS) questionnaire and the Youth Risk Behaviour Survey (YRBS; CDC, 2011, 2013). The questionnaire assessed children's weight, height, nutrition and PA behaviours, as well as their knowledge about experts' recommendations in the last two domains.

Two items assessing knowledge about experts' recommendations for nutrition (KERN) and PA ("How many fruits and vegetables do you think experts recommend you should eat in a day?" and "How much physical activity do you think experts recommend that kids your age should get?") were taken from the YPANS (Zapata et al., 2008). Children's nutrition behaviour was assessed from their self-reported consumption frequency of fruits, fruit juice, vegetable and green salad, milk, milk type, breakfast, and refreshments. Vegetables and green salad consumption was assessed by their responses to the questions:

During the past seven days, how many times...

a) did you drink 100% fruit juice, such as orange, apple, or grape juice? (Do not count sport drinks, or other fruit-flavoured drinks).

b) did you eat fruit? (Do not count fruit juice).

c) did you eat green salad?

d) did you eat vegetables? (Vegetables include foods like broccoli, spinach, carrots, squash, tomatoes or green beans).

Milk consumption was assessed using the questions:

a) During the past seven days, how many glasses of milk did you drink? (Include milk you drank in a glass or cup, from a carton or with cereal. A milkshake counts as a glass of milk. Count the half pint of milk served at school as equal to one glass).

b) When you drink milk, what kind of milk do you usually drink? Soda water consumption was measured by the question "During the past seven days, how many sodas did you drink?"

Breakfast consumption was assessed by the question "During the past seven days, how many mornings did you eat breakfast?"

Children's PA behaviour was assessed using modified items of the YPANS and YRBSS, in order to reflect the updated Healthy People 2020 objectives for children's PA (Centre of Disease Control and prevention [CDC], 2011; Zapata et al., 2008). The following items measured amount of PA:

a) On how many of the past seven days did you exercise or participate in physical activity for at least 60 minutes that made you sweat and breathe hard, such as basketball, soccer, running, swimming laps, fast cycling, fast dancing, or similar aerobic activities?

b) On how many of the past seven days did you exercise or participate in physical activity that *did not* make you sweat and breathe hard, such as slow cycling, skating, pushing a lawn mower, or mopping floors?

Parents' questionnaire.

The second questionnaire was similar to that for children. It contained the same items for demographics, weight, height, nutrition and PA behaviour, as well as knowledge of experts' recommendations in the last two domains. Parents' weight status was assessed and interpreted using the BMI according to the CDC guidelines (i.e., weight/height^2; CDC, 2010). Parents' PA behaviour was assessed using the items from the children's questionnaire, slightly modified to reflect the updated Healthy People 2020 objectives for adults PA (CDC, 2011; Zapata et al., 2008). More precisely, the questions that assessed amount of PA were:

a) How many hours in the past seven days did you exercise or participate in physical activity that *did not* make you sweat and breathe hard, such as fast walking, slow bicycling, skating, pushing a lawn mower, or mopping floors?

b) How many hours in the past seven days did you exercise or participate in physical activity that made you sweat and breathe hard, such as swimming laps, fast bicycling, fast dancing, or similar aerobic activities?

c) How many times on the past seven days did you participate in muscle strengthening activities?

Procedure

All items of the questionnaire were translated and back translated by two independent bilingual psychologists, as recommended by Hambleton (2001). An English language instructor examined the translations for clarity and readability and offered recommendations for minor suggestions in the wording.

Teachers to public secondary education school students administered the children's questionnaires to a convenience sample. Students were instructed by their teachers to complete the questionnaires after class and to ask one of their guardians at home to complete the parents' questionnaire. As an incentive, children were told that the completion of questionnaires would give them the right to enter a lottery and win a free entry participation in an athletic event (3 on 3 basketball tournament).

Teachers informed participants about the aim of the study and ensured about the confidentiality of their responses. Parents signed consent forms for them and on behalf of their children. From the initial pool of 200 children, 153 returned fully completed questionnaires (77% response rate). The completion of each questionnaire required approximately 5-8 minutes.

3   Results

Children's weight status was calculated and interpreted according to their Body Mass Index (BMI) percentile on the CDC BMI growth chart (CDC, 2011). Descriptive data showed that 80% of the children had normal weight, 4% were underweight, and 16% were overweight (from which 5% were obese). Among their parents, 52% had a normal weight, and 48% were overweight (from which 10% were obese). Pearson correlation analysis (two-tailed) was used to examine the relation of parents and children's BMI, in the diet and PA behaviours described above (see children's and patents' questionnaire) as well as in knowledge about experts' recommendation for nutrition (KERN) and PA. The significant correlations in various diet domains, KERN, PA and BMI between parents and children are presented in Table 1.

Chi-square analysis revealed that the preference for milk type was statistically significant for the pairs parents-children, $N = 129$, $\chi^2(6) = 31.24$, $p = .00$, mothers and children, $N = 89$, $\chi^2(4) = 21.01$, $p = .00$, mothers-sons, $N = 34$, $\chi^2(6) = 14.51$, $p = .02$, and mothers-daughters, $N = 62$. $\chi^2(6) = 14.51$, $p = .02$. More precisely, children who drank full fat milk had parents who drank also the same type of milk. Children who had mothers that drank low fat milk were more likely to drink low fat milk too. Sons who drank either full or low fat milk had mothers that drank the same type of milk respectively.

A stepwise regression analysis was conducted in order to identify predictors of children's BMI. The domains included in the analysis where those correlated between parents and children in the correlation analyses (i.e., BMI, KERN, fruit juice

Table 1

BMI, KERN, consumption frequency in various diet behaviour domains, and Physical Activity for Children and Parent, as well as correlations between parents and children in specific domains.

	Sons	Daughters	Mothers	Fathers	PC	MC	FC	MD	FD	MS	FS
	M	M	M	M							
BMI	19.8 (4.55)	18.96 (2.93)	24.14 (3.40)	26.68 (3.61)	.26**	.33**		.42**			
KERN	3.12 (1.63)	3.18 (1.35)	3.08 (1.12)	2.97 (1.54)	.39**	.36**		.37**			
Moderate PA	5.41 (2.52)	2.40 (2.51)	3.66 (3.03)	3.36 (3.31)					.81**		
Diet											
Fruits	6.35 (4.21)	5.09 (4.50)	5.26 (3.01)	5.36 (3.42)				.47**			
Fruits juice	3.58 (2.42)	3.61 (2.60)	2.43 (2.39)	2.29 (1.94)	.41**	.44**		.46**			
Vegetables	4.62 (3.31)	4.18 (4.00)	4.91 (1.89)	4.76 (2.63)				.52**		.52**	
Breakfast	6.08 (2.4)	5.41 (2.52)	5.65 (2.34)	5.24 (2.42)	.33**		.60**				
Green salad	3.96 (3.16)	3.4 (2.40)	4.30 (2.45)	4.98 (2.01)	.24**						.67**

Note: PA = Physical Activity, KERN = Knowledge about Experts Recommendations for Nutrition, BMI = Body Mass Index, PC = parents and children, MC = mothers and children, FC = fathers and children, MD = mothers and daughters, FD = fathers and daughters, MS = mothers and sons, FS = fathers and sons, numbers in parenthesis represent standard deviations

** $p < .007$ Bonferroni correction applied.

breakfast, and green salad consumption). In the first step, children's and parents' gender and age were included. In the second step of the analysis, parents' BMI, KERN, milk preference, as well as fruit juice, breakfast, and green salad consumption were included. At the third step, children's' KERN, milk preference, as well as fruit juice, breakfast, and green salad consumption were included. The overall model explained (AdjR^2) 17.9% of the variance in children's BMI [$F(15, 69) = 2.222$, $p = .013$]. Children's age ($\beta = .363$, $p = .005$), parents' BMI ($\beta = .286$, $p = .010$), and parents' breakfast consumption ($\beta = -.244$, $p = .049$), were significant predictors of children's BMI.

We repeated the same analysis using the same steps including the same domains for the subsample of mothers. The overall model explained (AdjR^2) 27.4% of the variance in children's BMI [$F(14, 41) = 2.481$, $p = .012$]. In the third step of the analysis, children's age ($\beta = .463$, $p = .004$), children's gender ($\beta = -.287$, $p = .028$), and mothers' BMI ($\beta = .344$, $p = .007$) were the significant predictors of children's BMI. We repeated the same analysis with the same steps including the same domains for the subsample of fathers, however no significant model emerged.

4   Discussion

This study was set out in Greece to examine if parental BMI, diets, PA behaviour, as well as the corresponding knowledge for optimal behaviour in diet and PA are related to those of their children. Furthermore, it was examined if parental diet and PA behaviour, and knowledge for optimal behaviour in these areas, are related to their children's weight status. The preliminary analysis of the data revealed the majority of children to have a normal weight status. Of them, 16% were overweight and 5% were obese, in contrast to 23% and 11%, respectively, in the Chaniotis et al. (2010) study. In extend, the children-participants in our study may not have been a representative sample. This may be caused by the incentives of partaking in our study (i.e., tickets in an athletic event). This type of incentive may have attracted more athletic children than non-athletic ones. As a consequence, the data may have been tainted by a sampling error of employing children that are generally more interested in doing sports, hence probably with less weight in average than sport-inactive children.

Despite that, the data revealed that children's BMI was positively correlated with their parents' BMI. This evidence supports our third hypothesis and is consistent with past findings by Malinandros et al. (2009) and Kosti et al. (2008). In the present study, maternal BMI was predictive of children's BMI, especially for their daughters. These results indicate that mothers' more than fathers weight status may be influential in shaping their children's weight status, especially for their daughters. Supporting our first hypothesis, this study indicated parents and children's diets to be related. This is in line with previous findings for parents and children's consumption of

healthy snacks (Brown & Odgen, 2004), unhealthy snacks (Tsamita & Kasteroliotis, 2008) and quality of diet (Lazaroy, Kalavana, & Matalas, 2008). From all nutrition categories examined, the analysis yielded positive correlations for fruit juice, milk type, green salad and breakfast consumption. As frequently parents ate items of these categories, as frequently did so their children too. In addition children tended to consume the same type of milk as their parents. Parent's breakfast consumption predicted their children's weight status in a positive manner; as more frequently parents ate breakfast, as less was the weight status of their children.

The relation between parents and children's diets were not present in a uniform manner. By examining the subsamples of mothers and fathers separately, mothers' diet was related to their children's in one nutritional category (i.e., fruit juice) and with fathers' in another (i.e., breakfast). Mothers' diet was correlated moderately with their daughters in 3 nutrition categories, but only in one with their sons. No such correlation was evident between fathers and their daughters, whereas with their sons, it was present only in breakfast consumption.

However, not mothers or fathers' diet were statistically predictive factors of children's BMI in a uniform manner. These finding do not support our third hypothesis, and are in odds with the arguments of Skinner et al. (2002), and Koui and Jago (2008), who pinpointed mothers' diet and preference in food as bordering factor of their children's diet. The data also fail to give full support to the arguments of Horn and Horn (2007), Golan and Crow (2004) and ADAF (2011), in respect to the gender-specific parental-modelling importance in nutrition.

Previous studies illustrated parents and children's PA to be related (Bois et al., 2005; Fogelholm et al., 2005). The present study did not detect similar trends, failing to support fully our first hypothesis. Data showed that only fathers' moderate PA was positively correlated with those of their daughters, a finding that has been left unexplained.

Knowledge regarding experts' recommendation for nutrition (KERN) was positively correlated amongst the members of a family, with the exception of the knowledge relation between fathers and their sons. Contrarily, there were no correlations evident amongst family members regarding knowledge of experts' recommendations for PA. Thus, the evidence argues in favour of the existence of a family systemic influence regarding knowledge about what is considered to be an optimal diet, supporting our second hypothesis partially. Yet, KERN failed to predict children's BMI in the whole sample, as well as by examining the mothers and fathers' subgroup separately. In extend, the evidence give only partial support to the arguments of Pratt (1976) and Ornela et al. (2007) for the existence of a family system regarding nutrition.

To further mud the water, children's age was positively related to their own BMI status: this may indicate that as children grow, their BMI status increases. Since BMI is a value that is corrected for age and height, this finding may surprise. A possible

explanation is that children tend to gain disproportionally more weight as expected, as they grow older. This may be cause having more time away from the family's nest, more independence and more opportunities to eat "prohibited food", children may eat and drink their way towards an unhealthier level of BMI status as they age.

An important non-finding of this study is the total failure to capture children's nutrition and PA behaviour as significant predictors of their own BMI status. This is surprising, since it is widely acknowledged that PA, fruits and vegetables, green salad, breakfast and the restriction of sodas are key components in weight control (Zapata et al., 2008). The facts that the children employed in this study may have been sports-oriented, their fairly high means in consumption of fruits and breakfast, and the restricted range of food categories surveyed in the questionnaires may offer a possible explanation. More precise, if most children have been sports-oriented, than it may explain why they have been in great extend within the range of "normal" weight.

Another explanation might be that the self-report, questionnaire based nature of this study might have failed to capture genuinely the participants' BMI, diet and PA behaviours. The consumption of important food categories such as sweets, meat, fish and whole grain products, which are equally important diet domains for weigh control, was left out of scope in our study. In contrast, national surveys that employ large samples of participants and survey more food categories (studies from which the importance of fruits and vegetables are pinpointed), are able to examine data of broader range than the present study, thus offering more significant correlations between these domains.

The same explanation may apply for PA. The fairly high means of moderate PA (i.e., 5.41 times per week) gives support to the assumption that the sample employed may have been sports-oriented. It could also further explain why children's subsample falls in great extends into the "normal" weight range of the BMI spectrum. As a result, the analyses in this study could only detect the fluctuation of children's BMI in a narrower range of weight, food categories, as well as consumption and PA frequency.

Further limitations of the study should be taken into account. The examination of nutrition and PA inside the family was based in a cross-sectional design. The pairs formed from the study's data were unequal (i.e., daughter-mother 75, son-father 20, daughter-father 22, son-mother 34). The study's cross-sectional design did not allow the examination of the influence between mothers and fathers, nor between sisters and brothers inside the family. Lastly, findings were based in correlation analyses. Correlation-design studies cannot demonstrate causation or direction, or the underlying mechanisms of a relation. Future research in Greece could address these issues by employing a larger sample, surveying more food categories in a longitudinal study.

.

In sum, the evidence of this study partially argues in favour of a family system regarding health behaviours as described by Pratt (1976) and Ornela et al. (2007). Weight, diet behaviours and nutrition knowledge was in great extend positively corre-

lated inside the Greek family. These relations were more evident between children and their mothers, rather between them and their fathers. Maternal BMI may be predictive of children's BMI, especially for their daughters, pinpointing a maternal influence in children's weight shaping. In total, the evidence indicates that parental influence in nutrition may not be directed only to the same sex offspring. It places noteworthy however, that healthy nutrition and a healthy weight is a "family issue" more than a personal one in young age.

5   References

American Dietetic Association Foundation. (2011). *The state of family nutrition and physical activity: Are we making progress?* Retrieved from http://www.eatright.org/foundation/fnpa/

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Benton, D. (2004). Role of parents in the determination of the food preferences of children and the development of obesity. *International Journal of Obesity, 28*, 858-869. doi:10.1038/sj.ijo.0802532.

Breen, F. M., Plomin, R., & Wardle, J. (2006). Heritability of food preferences in young children. *Physiology and Behavior, 88*(4-5), 443-447. doi:0.1016/j.physbeh.2006.04.016.

Birch, L. L., & Fisher, J. O. (1998). Development of eating behaviors among children and adolescents. *Pediatrics*, 101, 539–49.

Bois, J. E., Sarrazin, P. G., Brustard, R.J., Troulloud, D. O., & Cury, F. (2005). Elementary schoolchildren perceived competence and physical activity involvement: the influence of parents' role modeling behaviors and perceptions of their child's competence. *Psychology of Sport and Exercise*, 6(4), 381-397. doi: 10.1016/j.psychsport.2004.03.003.

Brown, R., & Ogden, J. (2004). Children's eating attitudes and behavior: a study of the modelling and control theories of parental influence. *Health Education Research, 19*, 261-271. doi: 10.1093/her/cyg040.

Center of Disease Control and Prevention, (2010). 2010 National Youth Physical Activity and Nutrition Survey. Retrieved from: http://www.cdc.gov/healthyYouth/yrbs/pdf/nypans/2010nypans_questionnaire.pdf

Center of Disease Control and Prevention, (2010). Youth Risk Behavior Surveillance-United States, 2011. Retrieved from: http://www.cdc.gov/mmwr/pdf/ss/ss6104.pdf

Chaniotis, D.I, Botsari, S., Mikelopoulou, P., & Chaniotis, F. (2010). Eating habits and obesity indices assessment in School Health based survey in the Attica basin Greek pupils: 2002-2007**.** *Ejournal of Science and Technology, 5*(3), 49-61. Retrieved from http://e-jst.teiath.gr/issue_16/Chaniotis_16.pdf.

Child obesity doubles in decade (2006, 21 April). BBC news. Retrieved from http://news.bbc.co.uk/2/hi/health/4930264.stm

Children's Obesity Facts. (2011, September 15). Center for Disease Control and Prevention. Retrieved from http://www.cdc.gov/healthyyouth/obesity/facts.html

Chu, S. Y., Grossman, M., & Saffer, H. (2004). An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System. *Journal of Health Economics, 23*(3), 565-587.

Cutler, D. M., Glaeser, E. L., & Shapiro, J. M. (2003). Why have Americans become more obese? *Journal of Economic Perspectives, 17*(3), 93-118. Retrieved from http://faculty.chicagobooth.edu/jesse.shapiro/research/obesity.pdf

Doak, C. M., Visscher, T. L. S., Renders, C. M., & Seidell, J. C. (2006). The prevention of overweight and obesity in children and adolescents: a review of interventions and programme. *Obesity Reviews, 7*(1), 111-136. doi: 0.1111/j.1467-789X.2006.00234.x.

Fogelholm, M., Nuutinen, O., Pasanen, M., Myohanen, E., & Saatela, T. (1999). Parent child relationship of physical activity patterns and obesity. *International Journal of Obesity and Metabolism Disorders, 23*(12), 1262-1268.

Gleitman, H., Reisberg, D., & Gross, J. (2007). *Introduction to psychology (7th ed.).* NY: W.W. Norton & Company.

Golan, M., & Crow, S. (2004). Targeting parents exclusively in the treatment of childhood obesity: long-term results. *Obesity, 12*, 357-361. Retrieved from http://www.shahaf-ltd.co.il/n/read_write/long_term_results_obes_res_2004.pdf

Gray, W. N., Kahhan, N. A., & Janicke, D. M. (2009). Peer victimization and paediatric obesity: A review of the literature. *Psychology in the Schools, 46*(8), 720-727. doi: 10.1002/pits.20410.

Hambleton, R. K. (2001). The next generation of the ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment, 17*(3), 164-172.

Horn, T. S., & Horn, J. L. (2007). Family influences on children's sport and physical activity participation, behavior, and psychosocial responses. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 685-711). Hoboken, NJ: John Wiley and Sons.

Janssen, I., Craig, W. M., Boyce, W. F., & Pickett, W. (2004). Associations between overweight and obesity with bullying behaviors in school aged children. *Pediatrics, 113*(5), 1187-1194.

Johnson S.L., & Birch L.L. (1994). Parents' and children's adiposity and eating style. *Pediatrics, 94*, 653-661.

Koplan, J. P., Liverman, C. T., & Kraak, V. A. [Ed.] (2005). *Preventing childhood obesity: Health in balance.* Washington, DC: The National Academies Press.

642

Kral, T., & Faith, M. (2009). Influences on child eating and weight development from a behavioral genetics perspective. *Journal of Pediatric Psychology, 34*(6), 596-605. doi: 10.1093/jpepsy/jsn037.

Kosti, R. I., Panagiotakos, D. B., Tountas, Y., Mihas, C. C., Alevizos, A., Mariolis-Sapsakos, et al. (2008). Parental Body Mass Index in association with the prevalence of overweight / obesity among adolescents in Greece; dietary and lifestyle habits in the context of the family environment: The Vyronas study. *Appetite, 51*(1), 218-222. doi: 10.1016/j.appet.2008.02.001

Koui, E., & Jago, R. (2008). Associations between self-reported fruit and vegetable consumption and home availability of fruit and vegetables among Greek primary school children. *Public Health Nutrition, 11*(11), 1142-1148. doi:10.1017/S1368980007001553.

Kring, A. M., Johnson, L. S., Davison, C. G., & Neale, M. J. (2010). Abnormal Psychology (7th ed.). Asia: John Wiley & Sons.

Lazarou, C., Kalavana, T., & Matalas, A. L. (2008). The influence of parents' dietary beliefs and behaviours, on children's dietary beliefs and behaviours: The CYKIDS study. *Appetite Journal, 51*(3), 690-696.

Lobstein, M. T., & Frelut, M. L. (2003). Prevalence of overweight among children in Europe. *Obesity Reviews, 4*(4), 195-200. Retrieved from http://easo.org/docs/report70.pdf

Malindretos, P., Doumpali, E., Mouselimi, M., Papamichail,N., Doumpali,C., Sianaba,O., et al. (2009). Childhood and parental obesity in the poorest district of Greece. *Hippokratia, 13*(1), 46–49.

Ornelas, I. J., Perreira, K. M., & Ayala, G. X. (2007). Parental influences on adolescent physical activity: A longitudinal study. *International Journal of Behavioral Nutrition and Physical Activity, 4*(3). Retrieved from http://www.ijbnpa.org/content/4/1/3

Pratt, L. (1976). Family structure and effective health behavior. The energized family. Boston: Houghton-Mifflin.

Razakou, F., Tsapakidou, A., Besi, C., Tsompanaki, T. (2003). Investigating concrete factors related to the 7-12 old children's occupation with out of school athleticism. Retrieved from http://www.hape.gr/emag/vol1_2/hape23.pdf

Risvas, G., Panagiotakos D. B., & Zampelas, A. (2008). Factors affecting food choice in Greek primary-school students: ELPYDES study. *Public Health Nutrition, 11*(6), 639 646. doi: 10.1017/S1368980007001073

Strauss, R. S., & Pollack, H. A. (2003). Social marginalization of overweight children. *Archives of Pediatrics and Adolescent Medicine, 157*, 746-752. doi:10.1001/archpedi.157.8.746.

Sarafino, P. E. (2008). *Health Psychology. Biopsychosocial Interactions (6th ed.).* Hoboken, NJ: John Willey & Sons.

Skinner, J. D., Carruth, B. R., Bounds, B., & Ziegler, P. J. (2002). Children's food preferences: A longitudinal analysis. *Journal of the American Dietetic Association, 102*, 1638-1647.

Sturm, R. (2002). The effects of obesity, smoking, and drinking on medical problems and costs. *Health Affairs*, 245-253. doi: 10.1377/hlthaff.21.2.245.

Tsamita, I., & Karteroliotis, K. (2008). Health related behaviours, social factors, and adolescents' dietary habits in a rural city in Greece. *Hellenic Academy of Physical Education, 6*(1), 25-36.

Wang, G., & Dietz, W. H. (2002). Economic burden of obesity in youths aged 6 to 17 years: 1979-1999. *Paediatrics, 109*(5), 1-6.

Welk, G. J. Wood, K., & Morss, G. (2003). Parental influence on physical activity in children: An exploration of potential mechanisms. *Paediatric Exercise Science, 15,* 19-33. Retrieved from: http://extranet.nuorisuomi.fi/download/attachments/3245041/Parental%20influences%20on%20physical%20activity%20in%20children.%20an%20exploration%20of%20potential%20mechanisms.pdf

Zapata, L. B., Bryant, C. A., Macdermott, R. J., & Hefelfinger, J. A. (2008). Dietary and physical activity behaviours of middle school youth: The youth physical activity and nutrition survey. *Journal of School Health, 78*(1), 9-18. Retrieved from: https://middleschoolobesity.wikispaces.com/file/view/Article.pdf

Qualified majority voting vis-a-vis consensus in the EU's Council of Ministers

Todorovic Jelena[1] and Grmusa Adrijana[2]

[1] Institute for political studies, Belgrade
todorovic.j82@gmail.com
[2] Institute for political studies, Belgrade
adrijana.grmusa@gmail.com

Abstract. In this paper the authors will analyze the evolution of decision-making process within the Council of Ministers. The analysis will focus on two dominant models of decision-making which are applied within this EU institution-qualified majority voting and decision-making by consensus. The evolution of the qualified majority voting involves constant redistribution of votes and its change. On the other hand, decision making by consensus is informal decision-making rule which gain its importance during the years of practice. Having in mind that Council of Ministers is central EU decision-making institution, the aim of this paper is to explain the importance of formal and informal decision-making rules in the Council of Ministers by using comparative method and qualitative method of social science.

Keywords: *Council of Ministers, Qualified Majority Voting, Consensus*

1   Introduction

The Council of Ministers is the central decision-making institution in the European Union (Nugent, 1999). By its nature-an intergovernmental body (body consisting of ministers of Member States), the founding treaties gave him the main place in the EU institutional architecture. The Council of Ministers is particularly interesting because of its hybrid, dual nature that distinguishes it from similar institutions of other international organizations. The hybrid nature of the Council is reflected in the following: a) Council has both legislative and executive features b) also has intergovernmental and supranational characteristics (Hix, 2005).

Speaking of supranational characteristics of the Council of Ministers, they are primarily evident in decision-making. Decision-making in the Council of Ministers may be unanimous, by a simple majority and qualified majority voting. When decisions are taken by a qualified majority, each Member State has a different number of (weighted) votes. Decisions taken by qualified majority should be applied even by the countries that did not vote for them (Renshaw-Hayes and Wallace, 1997). It reflects the crucial specificity of decision-making in the Council of Ministers.

Founding treaties designed different types of decision-making procedures depending on the area to be addressed in the Council of Ministers. When the Council of Ministers decides unanimously, or by a simple majority, then it shall apply the rule one state-one vote, the same principle as in other typical international organizations (UN, OSCE, NATO). Qualified majority voting was envisaged by the Treaty of Rome but it should have been applied only after the establishment of the Customs Union in the areas related to the completition of common market. During the years, the Council of Ministers has developed informal norms that have an increasing impact on decision-making-decision-making by consensus (Heisenberg, 2005).

Competitiveness between intergovernmental and supranational features within the Council of Ministers and the political struggle for supremacy between large and small Member States, attracted attention of authors and opted us to tackle the details of decision-making within this institution. The first struggle is reflected in the choice of decision-making model - unanimity or qualified majority voting, the other is reflected through decades of Member States pursuit to maximize their impact on EU decision-making.

2   Methodology

The theoretical starting point of this paper is closer to the liberal theory of intergovernmental cooperation. Intergovernmental cooperation theory argues that the key of the European integration is in the hands of the Member States that are integrated in so far it suits their national interests (Moravcsik, 2009). In this context, we observe the change of qualified majority model, as well as the expansion of qualified majority voting in the Council of Ministers to the extent and in those areas where it suits the Member States.

In this paper primarily will be used comparative method and qualitative method of social science. The comparative method will be applied as comparisons of formal decision-making rules that are changed with almost every revision of the founding treaties. Authors observed similarities and differences between proposed solutions and the evolution qualified majority voting. Qualitative methods of the social sciences will be applied through the analysis of books, articles and papers related to the subject. It is expected that this paper contribute to explanation of decision-making process within the Council of Ministers.

3   Literature review

Theoretical considerations and the importance of the role that the Council of Ministers has in the institutional structure of the Union illustrates the wider approach to the theory of European integration. Functionalist theory holds that the Council is the brake of deeper integration and to puts narrow national interests above the common European interest (Rosamond, 2000). Liberal theory of intergovernmental cooperation in turn, stands on the second position that it is good that member states hold strings of

integration in their hands, and so can determine which areas will be further integrated (Moravcsik, 2009).

The Council is not only the central decision-making institutions, but also the main forum for negotiations. The Council has developed a special model of negotiation (bargaining) which became a separate topic for research in the social sciences. A large number of scientific studies adresses the behavior of states in the 'bargaining' which is based on the application of game theory (Renshaw-Hayes and Wallace, 1997).

Decision-making in the European Union seems quite complicated, not only in terms that each new revision of the founding treaties has changed it but also that there is a huge number of different procedures that are applied depending on domains where the legislative proposal under discussion belongs. Most authors in the analysis of decision-making in the European Union mainly study the formal decision-making (Wallace and Wallace, 2000; Sherrington, 2000; Hix, 2005; Bomberg and Peterson, 1999).

This is logical considering because the decision-making in the Council of Ministers was characterized by a 'veil of secrecy' for almost four decades. 'Veil of secrecy' in this case meens that there was no release of Council voting data up to 1993. Since 1994, Council of Ministers is directed towards increasing the transparency of its work. Since that time the data on decision-making in the Council of Ministers are publicly available (Heisenberg, 2005). Authors who were engaged in research of these data (Renshaw-Hayes, Van Aken and Wallace, 2005; Mikko Mattila, 2004; Mattila, and Lane, 2001; Hagemann and Clerck-Sachsse, 2007; Heisenberg, 2005), came to the conclusion that consensus is actually the main form of decision-making and voting is rare, even when it is envisaged.

4   Decision-making by consensus

In 1965, in protest against majority voting, France mounted a boycott on all Community business, lifting it only when the national right of veto was restored by the 'Luxembourg Compromise' (Sherrington, 2000). This agreement stipulates that in case of 'very important interests' of a Member State, the discussion will lead to reaching consensus. Unanimity remained the dominant form of decision-making and thereby eliminate French fear that in the case of moving to qualified majority voting it can be outvoted. 'Luxembourg compromise' – explained by vital national interests, led to long negotiations that preceded the achievement of consensus. It seems that this is the cause of later decision-making by consensus without a formal voting even in those situations where it is envisaged to be decided by a qualified majority.

Therefore, based on data published by the Council between in 1994 and 2002, even 81% of the decisions was adopted by consensus without a formal vote (Heisenberg, 2005). There are several explanations. The first is based on the direct impact of the 'Luxembourg compromise'. For years after the 'Luxembourg compromise', Member States have had long negotiations in decisions-making, until all countries agreed with legislative proposal. Such practice has influenced that even after the introduction of qualified majority voting, Member States often made decisions without a formal vote.

Many countries do not want to be outvoted. If they repeatedly find themselves in the position of being a minoritywhich is not enough to block the decision, there is a danger for that the country to be isolated. Another explanation of such a use of informal institution of consensus is that if you make decisions based on consensus, it may provide greater stability of the Community. Such decisions will be easier to implement in practice, considering that the some of them reached broad agreement.

The key difference between consensus and qualified majority voting is that under consensus the minority needs to be appeased in some way, although not completely. It is possible that agreement on one piece of legislation then puts into play another, unrelated one (Heisenberg, 2005). 'A decision on one particular proposal may force a decision on another proposal, due to package deals or tactical maneuvering by a member state or the Presidency' (Sherrington, 2000).

5   Council of Ministers-a priori voting power and qualified majority voting

During the establishment of the Communities, population number was used as a criterion for the weighted votes. However, this criterion is not applied directly proportional which means that Member States have the votes in proportion to their population but the principle of degressive proportionality. It is a corrective principle that favors small states. They receive more votes than they should receive using the direct proportion.

It should be pointed that the number of weighted votes was not so important until the end of the eighties. Weighted votes gained in importance, along with the question what should be the threshold when Council of Ministers decides by qualified majority voting, just after the extension of qualified majority voting across the whole single market programme (Single European Act and the Maastricht Treaty). Since that time these questions have become central issues at the Intergovernmental Conferences.

The study of the Member States voting power is based on the correlation between the number of weighted votes of the Member States and the threshold needed for decision-making by qualified majority voting. The essence of this connection is in the impact of each Member State on decision-making in the Council of Ministers, bearing in mind the number of its weighted votes and the threshold for qualified majority. After the second, third and fourth EU enlargement, the growing number of authors dealt with a priori voting power of the Member States. This is esspecially noticed during the preparation of fifth enlargement (largest in the number of potential new Member States), it suddenly increased the number of articles dealing with this topic. The reason for this is in the simple explanation. As final result, almost every enlargement changed the number of votes of the Member States as well as the threshold for qualified majority voting, thus changing also a priori voting power (Pajala and Widgrén, 2004).

A priori voting power of the Member States is calculated by using several well-known methods such as various indices, including the famous Normalized Banzhaf Index (NBI). NBI actually indicates the probability that Member State with a certain

number of votes influence the outcome of the vote, in randomly selected legislative proposals that are subjected to qualified majority voting (Baldvin and Widgrén, 2004). The general assumption is that the probability of any winning coalition is equal. This is never the case in reality. Some Member States are more prone to certain coalitions than others, however, the methods used to calculate the voting power abstract the political and economic preferences, inclination to make coalitions (Hagemann and Clerck-Sachsse, 2007). All this represent remarks to such a positivist approach to measure voting power. The value of NBI ignores important facts such as whether the Member State is a net payer or receiver, which block belongs to-the Nordic or the Mediterranean, whether it is large, medium or small, old or new Member State. All these characteristics are important in the Council of Minister's decision-making.

Some authors believe that the most important criterion for the reform of the qualified majority voting in the Council of Ministers is to preserve the Union's ability to act (Baldvin and Widgrén, 2004; Widgrén, 2009). 'Passage probability' is value that indicates a high probability that the Council should adopt some decision on a randomly chosen issue. Assumption is that all states are equally willing to vote for or against a proposal (Baldvin and Widgrén, 2003). First, the computer calculates all possible coalitions among EU members, namely every possible combination of yes- and no-votes by EU members. The final value is obtained by dividing the number of winning coalitions by the number of all possible coalitions. This value should be taken with reserves because the conditions under which it is calculated are quite general; Member States in reality do not always have the same attitude towards the proposal under consideration. However, this value is important to study the efficiency of the complex voting system. The value of the 'passage probability' depends on three factors: the number of Member States, number of weighted votes and the thresholds required for qualified majority voting (Baldvin and Widgrén, 2004a).

Since 1996 and the Intergovernmental Conference held in Amsterdam, it was clear that the voting system in the Council of Ministers must be changed and adapted to the functioning of the EU25+. In Amsterdam, Member States have failed to agree on acceptable qualified majority voting system, but these issues are transferred to the next Intergovernmental Conference in Nice. Treaty of Nice has created a complex system of triple majority that failed to solve the problem of decision-making efficiency. Temporary solution from Nice was supposed to be replaced with the simpler and more efficient system proposed by the Convention. Neither this system of double majority suited to all Member States as there was a possibility that Poland and Spain dissatisfied with proposed voting system, block the signing of the the Treaty establishing a Constitution for Europe .

Convention which preceded the Treaty establishing a Constitution for Europe, took into account that both legitimacy shoud be respected- legitimacy of Member States and the legitimacy of citizens, as it has been repeatedly pointed out that the EU represents Union of countries and Union of citizens (Kiljunen, 2004). This 'dilemma of legitimacy' was at the center of the debate about the choice of voting model in the Council of Ministers. From 1958 to 2004 it was easy to reconcile these two legitimacy because it was small number of Member States and in addition, the ratio between

large and small states was 5-10. Now the structure of the Union has significantly changed with 10 countries that became members in 2004, only one can be categorized as large (Poland), while the number of small states doubled. Bearing in mind the enormous differences in population among Member States, it was difficult to arrange threshold on population that fits all. For example, six large countries in the EU27 have almost 70% of the total population of the Union, while the smallest 11 (states that have less than 5 million inhabitants) have only 5% of the population of the Union (Baldvin and Widgrén, 2004a). Speaking of the membership criterion, all states have equal share of 1/27 or 3.7%. For some Member States is drastic difference between the percentage of population and percentage of membership, such as Germany, where the ratio is 17.1% -3.7% or vice versa Malta 0.1% -3.7% (Baldvin and Widgrén, 2004).

One of the positive aspects of the Treaty establishing a Constitution for Europe is that qualified majority voting becomes the main form of decision-making, unless otherwise is stipulated (Kiljunen, 2004). According to Article 24 of the Treaty establishing a Constitution for Europe for the adoption of legislative decisions it is necessary to vote 55% of member states and 65% of the EU population. In order to constrain blocking power of the large states, Treaty envisaged that blocking minority must have, at least, four Member States. This restriction was imposed because of the Germany since Germany and two other large States could block some decisions. There is another reason for the introduction of this rule - coalition of three members who opposed the war in Iraq (Germany, France and Belgium) (Devuyst, 2004).

Last in a series of revisions of the founding treaties Lisbon Treaty does not differ much in terms of decision-making in the Council of Ministers comparing to the failed Treaty establishing a Constitution for Europe. Lisbon Treaty stipulates that a qualified majority on Commission's proposal from November 2014 would be at least 55% of Member States, including the votes 15 of them and 65% of the EU population, while blocking minority is,at least, four Member States (Griller and Ziller, 2008).

6   Conclusions

A qualified majority voting has received broader application over the years and according to the Treaty of Lisbon, decision-making by qualified majority becomes the main form of decision-making. Another innovation of the Treaty of Lisbon is the abolition of of weighted votes-all Member States will have the equal number of votes in the Council of Ministers (Griller and Ziller, 2008). Based on the increasing number of Member States-nearly five times more than it was during the formation of the Communities, in the future it will not be easy to predict coalition between Member States. The thresholds for qualified majority and blocking minority must be seriously taken in account as they determine whether the EU will more or less frequently deal with crisis.

The formal decision-making rules are set in manner to prevent minority veto. The purpose of decision-making reform in the Council of Ministers is to prevent potential blockage of some Member States in the case of lower threshold for qualified majority

voting. This is reflected in avoiding either the overweighting of the small countries' parochial interests that could destabilize the EU or avoiding the tyranny of the large countries (Heisenberg, 2005). The impact of formal rules (thresholds for qualified majority voting, blocking minorities) is critical in situations where countries like Poland or Spain could block a decision, which is important for the whole EU. Furthermore, in our opinion, the EU is primarily and essentially an economic union, and veto players with greater economic power determine the rules that go in their favor, both, in the Council and in other EU institutions. Therefore, when EU is faced with changing the formal decision-making rules, the interests of the economically powerful Member States, mainly Germany, inevitably are taken into account. It is unacceptable that country that is the biggest financer of the EU budget could be brought into the situation of being outvoted in the Council of Ministers. This conclusion supports Moravcsik's assumption that the driving forces of EU integration lie in the interests of Member States and in the power they bring to Brussels (Moravcsik, 1998).

One of the conclusions is that the history of decision-making in the Council of Ministers is actually a history of the many attempts to strike balance between large and small states, between the very large, as well as between 'old' and 'new' Member States. New members, unaccustomed to decision-making by consensus, tightly remain on their positions during the voting process thus demonstrating their voting power. The old Member States were aware of this, and in the Treaty of Nice, particularly in the Treaty establishing a Constitution for Europe, strongly fought to preserve greater voting power in relation to the new members to ensure the stability of the EU and prevent its stagnation caused by potential blockages. The struggle over the distribution of votes between large and small states is in fact struggle to protect the national interest of ones or the others.

In the case when qualified majority voting is envisaged, Member States often decide by consensus rather than by formal voting. In our opinion, this fact is likely consequence of the perfect form of intergovernmental negotiation in which dominates the rule 'give something to get something' (Germany will support the French proposal in the field of agriculture that it is of vital importance if France supported the German proposal for the completion of the internal market for example) than it is due to 'Luxembourg compromise'.

If decision-making by consensus is considered to be the norm, it is difficult to understand the Member States' taft debates about the reallocation of Council voting weights with every accession of new members or the reluctance of Member States to apply the use of QMV to new issue areas (Heisenberg, 2005). One of the possible explanation is that Member States tend to make consensus as long as it allow them to protect their autonomous preferences that are mainly economic (Moravcsik, 2009). Otherwise, the formal decision-making rules would be applied bearing in mind that these rules are set in manner to prevent minority veto either large or small EU Member States.

Acknowledgements

References

Baldvin,R., Widgrén, M. (2003), The draft constitutional treaties voting reform dilemma, CEPS Policy Brief, No.44, http://aei.pitt.edu/1869/1/PB44.pdf

Baldvin,R., Widgrén, M. (2004) Winners and losers under various dual majority voting rules for the EU's Council of Ministers, CEPS Policy Brief, No.50, http://aei.pitt.edu/6546/1/1116.pdf

Baldvin,R., Widgrén, M. (2004a) Council voting in the constitutional Treaty-Devil in the details, CEPS Policy Brief, No.53
Baldvin,R., Widgrén, M. (2005) The Impact of Turkey's membership on EU voting, CEPS Policy Brief, No. 62, http://www.ceps.eu/files/book/1194.pdf

Bomberg, E., Peterson, J. (1999) Decision-making in the European Union, Palgrave, Basingstoke

Devuyst, Y. (2004) EU decision making after the treaty establishing a constitution for Europe, Policy Paper No.9, European Union Center, www.ucis.pitt.edu/cwes

Griller, S., Ziller, J. (2008) The Lisbon Treaty-EU constitutionalism without a constitutional treaty?, Springer, Wien New York

Hagemann, S., Clerck-Sachsse, J. (2007) Decision-Making in the Enlarged Council of Ministers: Evaluating the Facts, CEPS Policy Brief, No.119, http://aei.pitt.edu/7360/2/7360.pdf

Heisenberg, D.(2005) The institution of consensus in the European Union: Formal versus informal decision making in the Council, European Journal of Political Research, No.44, 65-90
Hix, S. (2005) The Political system of the European Union, Mackmillan Press LTD, London

Hosli, M., Mattila, M., Uriot, M.(2011) Voting in the Council of the European Union after the 2004 Enlargement: A Comparison of Old and New Member States, Journal of Common Market Studies, Vol.49, No.6, 1249-1270

Kiljunen, K. (2004) The European Constitution in the making, Centre for European Policy Studies, Brussels

Mattila, M. (2004) Contested decisions: Empirical analysis of voting in the European Union Council of Ministers, European Journal of Political Research, Vol.43, No.1, 29-50

Mattila, M., Lane, E. (2001) Why unanimity in the Council? A roll call analysis of Council voting, European Union Politics, Vol.2, No.1, 31-52

McIntyre, E. (1954) Weighted Voting in International Organizations, International Organization, Vol. 8 No.4, 484-497

Moravcsik, A. (2009) Liberal Intergovernmentalism, in Wiener, A. & Diez, T., (eds.) European Integration Theory, Oxford University Press, Oxford, 67-87.

Moravcsik, A. (1998) The Choice for Europe: Social Purpose and State Power from Messina to Maastricht, Cornell University Press, Ithaca

Nugent, N. (1999) The Government and Politics of the European Union, Palgrave, London

Pajala, A., Widgrén, M. (2004) A priori versus empirical voting power in the EU Council of ministers. European Union Politics. Vol.5, No1, 73-97.

Plechanovova, B. (2003) The treaty of Nice and the distribution of votes in the Council–Voting power consequences for the EU after the oncoming enlargement, European Integration Online Papers (EIOP), Vol.7, No.6,   http://eiop.or.at/eiop/texte/2003-006a.htm.

Renshaw-Hayes, F., Wallace, W., Wallace, H.(2005) When and Why the Council of Minsters of the EU votes explicitly, EUI Working papers RSCAS, No.25, http://aei.pitt.edu/7084/1/05_25.pdf

Reshaw-Hayes, F., Wallace, H.(1997) The Council of Ministers, Mackmillan Press LTD, London

Rosamond, B. (2000) European Integration Theory, Oxford University Press, Oxford
Sherrington, P.(2000) The Council of Ministers: political authority in the European Union, Oxford University Press, Oxford

Tsebelis, G., Yataganas, X. (2002) Veto players and decision making in the EU after Nice:Policy stability and bureaucratic/judicial discretion, Journal of Common Market Studies, Vol.40, No.2, 283-307

Wallace, H., Wallace, W. (2000) Policy making in the EU, Oxford University Press, Oxford

Wiberg, M. (2005) New winners and old losers, A priori voting power in the EU 25, Discussion Paper No.149, http://aei.pitt.edu/6823/1/dp_c149_wiberg.pdf

Widgrén, M. (2009) The impact of Council voting rules on EU decision making, CE-Sifo Economic Studies, Vol. 55, No.1, 30-56

From Re–use of Public Sector Information to Open Data

Kalina Georgieva[1]

[1] Department of Theory and History of the Legal systems, Faculty of Law, University of Sofia
"St. Kliment Ohridski", Bulgaria
kalina.georgievaa@gmail.com

Abstract. In the new era of ICT modern states are the most powerful collectors and source of all kind of information. Governments often hold very rich data, much of which is made public, but it still rests a significant part that is hidden from the public. Therefore the main issue under investigation in this paper is the Public Sector In-formation. It refers to non-personal data produced, stored or collected by public sector bodies. The topic is based on several guidelines, such as economy of knowledge, the governmental transparency and economic growth. Once this in-formation is made publicly available it gives us the chance to discover new and innovative solution and boost sustainable development. On other hand, it fosters the participation of citizens in political and social life and increases the transparency of government. The aim of the study is to show that Public data has significant potential for re-use in new products and services. For this reason, it describes the findings in the new legislation of the PSI. With the new European le-gal framework concerning the reuse of public sector infor-mation, the Directive 2013/37, significant part of the barriers in front of the ac-cess to public sector in-formation are erased and it is created a genuine right to re-use of public sector in-formation. The research methods used are desk study, comparative analysis and case study.

Keywords: *Information society, Public sector information, Open government data, Public sector information re-use.*

1   Introduction

One of the implications of the development of the Internet has been to promote the economic value of information. Information as an immaterial resource that can easily be shared, combined and transformed, has given rise to a growing number of products and services, with the subsequent creation of economic activity.

Therefore the information and the regulation of the information consists major topic in the Europe 2020 Strategy. The central aim of the EU 2020 strategy is to put Europe's economies onto a high and sustainable growth path. To this end, Europe will have to strengthen its innovative potential and use its resources in the best possible way. Europe 2020 sets out a vision of Europe's social market economy for the 21st

century. One of the priority themes in that context is 'Smart growth: developing an economy based on knowledge and innovation'. One of these resources is public data 1 — all the information that public bodies in the European Union produce, collect or pay for. The Commission's work in the area of open data is focusing on generating value through re-use of a specific type of data – public sector information, sometimes also referred to as government data. Examples are geographical information, statistics, weather data, data from publicly funded research projects, and electronic books from libraries. This information has a significant — currently untapped — potential for re-use in new products and services and for efficiency gains in administrations. Overall economic gains from opening up this resource could amount to € 40 billion a year in the EU. Opening up public data will also foster the participation of citizens in political and social life and contribute to policy areas such as the environment.

2   Literature Review

The used sources are mainly legal acts, international and of the EU such as: the Convention on fundamental rights and liberties, the European Charter of human rights , the Aarhus Convention , Regulations (EC) No 1049/2001 , Regulation (EC) No 1367/2006 , Recommendations (81) 19 of the Council of Europe , Recommendation (2002) 2 of the Council of Europe , Decision 2006/291/EC , Decision 2011/833/EU , Convention on Access to official documents from 2008 , Directive 2003/98 , Directive 2013/37/ EU , Communication of the EC to the EP ant the Council, the EcoSoCo and the Committee of regions on the re-use of PSI, review of Directive 2003/98/ EC COM (2009) 212 final and many others formal and informal acts.

Also the Green book to public access to documents, held by the institutions of the European community from 2007, the Report from the Commission on the application in 2005 of Regulation (EC) No 1049/2001 regarding public access to European Parliament, Council and Commission documents, Opinion of Article 29 Working party N 06/2013 on open data and public sector information ('PSI') reuse, Adopted on 5 June 2013. For the practical side of the topic, a useful article is „Unlocking the Potential of Public Sector Information with Semantic Web Technology" by Harith Alani, David Dupplaw, John Sheridan, Kieron O'Hara, John Darlington, Nigel Shad-bolt, and Carol Tulio.

3   Definition of a re-use of public sector information and its importance

The text of Directive 2003/98 as the first legal framework on the reuse of PSI, says that "'reuse' means the use of documents by persons or legal entities of documents, for commercial or non-commercial purposes other than the initial purpose for which the documents were produced."

In addition Decision 833/2011 says that "the exchange of documents between the Commission and other public sector bodies which use these documents purely in the

pursuit of their public tasks does not constitute reuse". In other words, re-use of public sector information (PSI) means using it in new and innovative ways by adding value to it, combining it with information from different sources, making "mash-ups" and new applications, both for commercial and non-commercial purposes. The ICT have created unprecedented possibilities to aggregate and combine content from different sources. The usage of ICT is important to extend to form a separate European institution to manage it. This is the Digital agenda for Europe.

What information might be available for re-use?

As Article 17(1) of Regulation (EC) No 1049/2001 regarding public access to Europe-an Parliament, Council and Commission documents provides, each institution must publish an annual report stating the number of cases in which it has refused to grant access to documents and the reasons for such refusals.

The Directive mentions very broad topics such as economic information, and information about patents and businesses. Public authorities hold lots of other information which could have a commercial value. Examples include data about properties, land use, lists of contractors, geographic data, statistics and everything and web site content. For example, local authority websites often have searchable information about tourism and leisure attractions and events. This content could be licensed to internet based holiday and leisure operators.

The first PSI Directive is having a big impact on the public sector. Public sector organizations are being given the ability to exercise more power over their information, to recover some of the costs of producing large and complex databases of information, and to some extent, to develop new commercial products. The point of the Directive and the Regulations is to develop an information industry. What is almost certain, as far as public authorities are concerned, is that the PSI Directive is "the Next Big Thing!"

4   The Digital Agenda for Europe

The Digital Agenda (DA) for Europe is one of the European Union's seven flagships for achieving its ten-year growth strategy. It is managed by the European Commission's Directorate General for Communications Networks, Content & Technology (DG CONNNECT). The aims of the DA are to update EU Single Market rules for the digital era. First of the goals, set in the Digital agenda is Digital Single Market. The importance of the matter is underlined by the Commission because too many barriers still block the free flow of online services and entertainment across national borders. The Digital Agenda places under Action number 3: To open up public data resources for re-use.

Public sector information is an important source of potential growth of innovative online services through value-added products and services. Governments can stimulate content markets by making public sector information available on transparent, effective and non- discriminatory terms. For this reason, the Digital Agenda for Europe singled out the reuse of public sector information as one of the key areas for action.

It is also created the EPSI plus portal. The European Commission established a network known as EPSI plus to monitor progress and share best practice across the Community.

5   Legal framework- the PSI Directive

Directive 2003/98, known as the Re-use Directive, established a minimum set of legal rules governing the re-use of PSI at the European level. It regulates in a unique way the third party use of information in the same conditions for those interested in it.

The Directive 2003/98/EC of the European Parliament and of the Council on the re-use of public sector information (the Directive) was adopted on 17 November 2003. With this directive EU member states jointly promote the re-use of public sector information within the single market. The European Commission required member states to implement the Directive by 1 of July 2005. Eight Member States meet the Commission's deadline. By the end of June 2007, 22 Member States had implemented the PSI Directive.

The Directive is an attempt to remove barriers that hinder the re-use of public sector information throughout the Union. The Directive encourages EU member states to make as much public sector information available for re-use as possible. Previous to the creation of this directive this area was left to member states to regulate. Nevertheless, the Directive doesn't contain any obligations for member states to provide the information to the public.

The Directive applies to all public sector information. This is any information that is produced by a public sector body including central government departments, local authorities and the NHS. The Directive does not apply to: Public service broadcasters, Educational and research establishments, such as schools, universities, archives, libraries and research facilities, Cultural establishments, such as museums, orchestras, operas, ballets and theatres.

The first PSI Directive was economic. Public sector information is a valuable information resource that could be used by the private sector to develop value added products and services. Providing significant economic opportunities and enhancing job creation across Europe, the Directive is a stimulus to the European economy in general.

In particular, it aimed to provide for transparency and consistency in the re-use of PSI to ensure a level playing field for all re-users of such information. However, the rules applicable under the Re-use Directive and as transposed under the Re-use Regulations only apply where a public sector body permits re-use of information covered by that legislation. Consequently, many public sector bodies across the European Union have opted not to make all their PSI available for re-use. The Commission points, however, to a lack of awareness across public sector bodies and an inconsistency of approach across Member States that has hampered the creation of cross border information products and services. This latest proposal seeks to bring about a further degree of harmonization at European level.

6   The revision of the PSI Directive

The aim of the Commission is for the new Directive to remove the obstacles that stand in the way of re-using public sector information. In doing so, it will stimulate the development of value added products and services across Europe and help to boost the information industry.

The Commission's proposal is a package of measures designed to help overcome existing barriers, as part of the Digital Agenda for Europe. The package consists of three strands which the Commission states will reinforce each other:

•   adapting the legal framework for data re-use, including through a proposal for a revised Directive on the re-use of public sector information (the re-use Directive), a revised Commission Decision on the re-use of its own information and soft law and policy measures;

•   mobilizing financing instruments in support of open data in research, development, innovation, competitiveness programmers and European infrastructure programs; and

•   facilitating coordination and experience sharing across the Member States.

Article 13 of the Re-use Directive called for a review of the application of the Re-use Directive before 1 July 2008. This review was carried out by the Commission. The Commission concluded that, despite the progress made, a number of barriers remained, and that a further review should be carried out by 2012 when more evidence of the impact, effects and application of the Re-use Directive would be available. The Commission undertook this review throughout 2011 and concluded that the following problems remained: insufficient clarity and transparency of PSI re-use rules, restrictive licensing conditions, lack of information on available data, lack of a robust complaints procedure, locked information resources, excessive charging and lack of a level playing field, unfair competition practices, insufficient enforcement and inconsistent approaches adopted by individual Member States.

The key features of this proposal are that it:

•   changes the general principle of the Re-use Directive so that Member States would be required to ensure that all PSI that is generally available to the public, subject to some specified exceptions, must be available for re-use for commercial and non-commercial purposes;

•   provides that where charges are made for the re-use of PSI the amount that can be charged should be limited to the marginal costs of reproduction and dissemination, unless exceptionally justified by reference to objective, transparent and verifiable criteria. In such exceptional cases, public sector bodies would be al-lowed to charge more than marginal costs, in particular where public bodies generate a substantial part of their operating costs from the exploitation of their intellectual property rights. Where charges are made, total income would as now not exceed the cost of collection, production, reproduction and dissemination together with a reasonable return on investment;

•   where licenses are required to re-use information, standard terms and conditions should be offered;

•	requires public sector bodies to define their "public task" in law or other binding rules rather than also by reference to common administrative practice;

•	expands the scope of the Re-use Directive to include libraries, archives, museums and university libraries, subject to the limitation that those institutions may choose whether to permit re-use and may charge over and above the marginal costs for the re-use of documents they hold; and

•	requires Member States to provide that re-users of PSI may seek redress for non-compliance with the Re-use Directive from an independent authority vested with specific regulatory powers and whose decisions are binding on public sector bodies.

According to the press release right before the vote for the Open data directive, the European Commission welcomes endorsement by the EU Council's 'Coreper' committee (EU Committee of Member States' Permanent Representatives) of the Commission's effort to open-up public sector data for re-use across Europe.

# 7	Directive 2013/37 or the Open Data directive

The revised Directive was adopted by the EU Parliament on the 26th of June 2013. Open data refers to the idea that certain data should be freely available for use and re-use. The main changes introduced by the New Directive are:

•	create a genuine right to re-use public information, not present in the original 2003 Directive;

•	expand the reach of the Directive to include libraries, museums and archives;

•	establish that public sector bodies can charge at maximum the marginal cost for reproduction, provision and dissemination of the information. In exceptional cases, full cost recovery (plus a reasonable return on investment) will remain possible;

•	oblige public sector bodies to be more transparent about charging rules;

•	encourage the availability of data in open machine-readable formats

Once fully implemented into national law, the revision of the 2003 Public Sector Information Directive would make all generally accessible (that is, non-personal) public sector information available for re-use. Developers, programmers, businesses and citizens will be able to get and re-use public sector data at zero or very low cost in most cases. They will also have access to more exciting and inspirational content, for example including materials in national museums, libraries and archives.

European Commission Vice-President Neelie Kroes and responsible for the DAE said that: "Opening up public data means opening up business opportunities, creating jobs and building communities. I welcome the Council's agreement to this culture change."

# 8	Open portals

Under the regulation of Directive 2003/98 the Commission has set an example to public administrations in making statistics, publications and the full corpus of Union law freely available online. This was a good basis to make further progress in ensuring the availability and reusability of data held by the institution.

The adoption of the new Directive leads to the introduction of a portal providing access to Commission data and datasets from across the EU. According to the Commission Decision of 12 December 2011 on the reuse of Commission documents, "A data 'portal' refers to a single point of access to data from a variety of web sources. The sources generate both the data and the related metadata. The metadata needed for indexing are automatically harvested by the portal and integrated to the extent needed to support common functionalities such as search and linking. The portal may also cache data from the contributing sources in order to improve performance or provide additional functionalities."

The intention was to establish a separate pan-European portal that would be in place from 2013. The Commission envisaged that this portal would link to portals and sites across Member States' own data portals such as the UK's data.gov.uk. For now the website has two main parts Data and Applications.

However, the European Union Open Data Portal is the single point of access to a growing range of data from the institutions and other bodies of the European Union (EU). Data are free to use and re-use for commercial or non-commercial purposes.

By providing easy and free access to data, the portal aims to promote their innovative use and unleash their economic potential. It also aims to help foster the transparency and the accountability of the institutions and other bodies of the EU and to "enhance the image of openness and transparency of the Institutions"

9 Conclusion

The path that the European Commission took in 2003 with Directive 2003/98 or the Public sector information Directive has now moved to another level. The initial legal framework was a giant first step, because it gave the legal definition for the term "reuse", stated the leading principles that are in force today and fostered the development in the sector. But it has also showed to the EU has a lot more to do concerning its policy for the re-use of PSI. The efforts of the Commission to surmount what was missing in the first Directive were brought to the adoption to the Open data directive 10 years later. In 2013 we can deliberately say that we have the following: a vibrant digital single market, the information as its main good, a genuine right of re-use of PSI, a newly introduced principal of access to all public documents, which is mandatory for all public bodies to follow, provision of the PSI at a zero or marginal price and that all of that creates the information industry as part of the information society. What needs to be done from now on is to transpose in the appropriate way Directive 2013/37/EU throughout the measures the each member state will chose for itself. So that the concept of the open data is brought to life for private persons and legal entities in the EU member states. The commercial re-use of public sector information that the ICT technologies enable, opens up countless opportunities for the development of new information products and services, driving forwards and accelerating the development of the knowledge economy.

10 References

1. OJ C 364, 18.12.2000, p. 1, Charter of fundamental rights of the European union (2000/C 364/01);
2. Convention for the Protection of Human Rights and Fundamental Freedoms, Rome, 4.XI.1950;
3. Convention on access to information, public participation in decision-making and access to justice in environmental matters, done at Aarhus, Denmark, on 25 June 1998;
4. Convention on Access to official documents from 2008. Available at; https://wcd.coe.int/ViewDoc.jsp?id=1377737;
5. OJ L 145, 31.5.2001, p. 43, Regulation (EC) No 1049/2001 of the European Parliament and of the Council of 30 May 2001 regarding public access to European Parliament, Council and Commission documents;
6. OJ L 264, 25.06.2006, p. 13, Regulation (EC) No 1367/2006 on the application of the Aarhus Convention;
7. OJ L 8, 12.1.2001, p. 1, Regulation (EC) No 45/2001 of the European Parliament and of the Council of 18 December 2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data;
8. OJ L 87, 31.3.2009, p. 164., Regulation (EC) No 223/2009 of the European Parliament and of the Council;
9. OJ L 345, 31.12.2003, p. 90, Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information;
10. OJ L 175 26.06.2013 p.1, Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013amending Directive 2003/98/EC on the re-use of public sector information;
11. OJ L 330, 14.12.2011, p. 39, Commission Decision of 12 December 2011 on the reuse of Commission documents, (2011/833/EU);
12. OJ L 107, 20.4.2006, p. 38, Commission Decision of 7 April 2006 on the reuse of Commission information (2006/291/EC, Euratom);
13. OJ C 240, 24.07.2014, p. 1, European Commission Notice, Guidelines on recommended standard licences, datasets and charging for the reuse of documents, COM (2014/C 240/01);
14. COM(2007) 185 final, Green book to public access to documents, held by the institutions of the European community from 2007;
15. COM (2009) 212 final, Communication of the EC to the EP ant the Council, the EcoSoCo and the Committee of regions on the re-use of PSI, review of Directive 2003/98/ EC;
16. COM(2011) 882 final, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of regions, Open data An engine for innovation, growth and transparent governance, Brussels, 12.12.2011;

17. COM (2007) 548 final Report from the Commission on the application in 2005 of Regulation (EC) No 1049/2001 regarding public access to European Parliament, Council and Commission documents, Brussels, 24.9.2007;

18. COM(2011) 877 final - 2011/0430 (COD), Communication of EcoSoCo on the ptoposal for directive for review of Directive 2003/98/ EC on the re-use of PSI;

19. Recommendations (81) 19 of the Committee of Ministers of the Council of Europe, to member states on the access to information held by public authorities, Adopted on 25 November 1981, at the 340th meeting of the Ministers' Deputies;

20. Recommendation (2002) 2 of the Council of Europe;

21. Digital agenda for Europe; s Website.
Available at: http://ec.europa.eu/digital-agenda/en/our-goals/pillar-i-digital-single-market;

22. Website: "European Union opendata portal". Available at: http://open-data.europa.eu/en/data/;

23. The ePSIplatform of the European Commission. Available at: http://www.epsiplatform.eu/;

24. Opinion of Article 29 Working party N 06/2013 on open data and public sector information ('PSI') reuse, Adopted on 5 June 2013;

25. Website: Information law, by Ibrahim Hassan, Available at: http://www.informationlaw.org.uk/;

26. Re-use of Public Sector Information in Spain: The Solid Foundations of a Promising Sector. Available at:
http://administracionelectronica.gob.es/pae_Home/dms/pae_Home/documentos/OBSAE/pae_Notas_Tecnicas/2011-12_nota_tecnica_RISP_EN/December_2011_-_RISP.pdf;

27. European Commission Press release, Commission welcomes Member States' endorsement of EU Open Data rules, Brussels, 10 April 2013;

28. Website: "Open Definition". Available at: http://opendefinition.org/od/;

29. JOINT WPI-COV SEMINAR on "Open Government In The Making"– Council's premises REVISED Concept Paper -scope, definitions, objectives of the seminar, practical approach, Brussels, 12 September 2012;

30. Alani, H., Dupplaw, D., Sheridan, J., O'Hara, K., Darlington, J., Shad-bolt, N., Tulio, C. Unlocking the Potential of Public Sector Information with Semantic Web Technology. 2007, Available at: http://eprints.soton.ac.uk/264429/ and http://oro.open.ac.uk/20028/1/;

31. Zijlstra, T., & Janssen, K The new PSI directive - as good as it seems? 2013, April 18.

Positive Dyslexia: An investigation into the causes and effects of stress in the classroom

Panagiota Blouchou[1]

[1]Department of Psychology, University of Sheffield
pblouchou1@sheffield.ac.uk

Abstract: This paper is focused on the influences of math and reading anxiety experienced by students with dyslexia, as well as teachers' coping strategies in mainstream schools, on those students' school performance and their subjective experience of schooling. To date, most definitions and research of dyslexia has focused on weaknesses and failures rather on potential strengths and positive aspects of dyslexia. To build on positive traits of dyslexia though, we need to identify the barriers to students' learning and embrace effective interventions to solve the problems. It has been found that stress, reading and maths anxiety cause low-performance at school and produce negative feelings in the learning process (Jansen et al., 2013). Teachers' stress in the classroom is also a significant factor. The present study will investigate the causes and levels of stress among students' and teachers' performance in the classroom. This will be examined using a qualitative study design which consists of two parts; in the first part interviews will be conducted with primary school children, dyslexic and non-dyslexic, aged 9-11 years old to detect the levels of maths and reading anxiety. Later on, four interventions will be implemented to students, which have been proved very effective for enhancing students' literacy and numeracy skills. The second part will consist of interviews with primary school teachers to elicit information on the major stressors and teachers' coping mechanisms. The study builds on two previous relevant studies on dyslexic students' performance and teachers' perceptions of working with them. The first one was a qualitative study which was conducted as part of my MRes in the University of Sheffield, which focused on an investigation into the teachers with dynamics working with students with dyslexia. The second study builds on an undergraduate project in the department of Psychology in the University of Sheffield, which investigated the effects of stress on the performance of dyslexic and control students.

Keywords: positive dyslexia; stress; maths anxiety; reading anxiety

1.   Introduction

Dyslexia is the most common learning disability, neurobiological in origin, with a prevalence of 4-10% in the population. It is a disorder with lifelong effects for dyslexic individuals and it manifests itself differently throughout lifetime requiring constant readjustment and change (Bruck, 1989). The most traditional definition of dyslexia is "a disorder in children who, despite conventional classroom experience, fail to attain the language skills of reading, writing and spelling commensurate with their intellectual abilities" (World Federation of Neurology, 1968).

Positive dyslexia is a new movement which was introduced by Professor Rod Nicolson in 2013. It is a new concept which continues the research on dyslexia as a disability but focuses on the strengths of students with dyslexia and how these can be identified, accelerated and applied throughout an individual's life (Nicolson & Agahi, 2013).

In this paper, the causes of stress and how it affects children to perform well in literacy tasks will be discussed, as well as the major sources of high levels of teachers stress. The first part of my presentation will focus on the discussion of data from the two previous studies that constitute the basis of my research.

The aim of my first study was to evaluate the teachers' experience and difficulties of working with dyslexic students, as well as understanding the impact of dyslexia on students' behaviour. In – depth semi-structured interviews were conducted with teachers and the approach used to analyse the data was the IPA (Interpretative Phenomenological Analysis, Smith, 2008). The findings suggested that the major difficulties experienced by teachers are due to the lack of adequate training, knowledge and expertise in the field of dyslexia. The fact that they were mainstream teachers made them feel stressed and anxious as there was only limited provision for teaching and learning strategies and IEPs (Individualised Educational Plans), likewise inadequate curriculum accommodations. Teachers' ratings of their dyslexic students' behavioural and emotional problems showed that they constantly refer to themselves in a maladaptive way, which consequently affects their self-concept and self-esteem. It was found that they mainly felt frustrated, anxious and ashamed when they had to read or write in the presence of their classmates or even bullied and disappointed by their peers.

The second study aimed at assessing the effects of stress on the performance of dyslexic and control primary school children on a range of tasks designed to investigate the declarative and procedural learning. This study was a two part investigation; the first part explored whether dyslexic children perform worse when put under stress, rather than when they have no pressure and no constraints. The

second part was an interview regarding the strengths and interests of dyslexic children compared to controls, in line with 'The Big Six Skills' (Big picture thinking, visuo-spatial processing, creativity, empathy, communication and teamwork) identified by Agahi & Nicolson (2013). This was a mixed measures qualitative and quantitative design. In the experimental condition students had to perform six different tasks (one minute reading, rapid automatized naming, motor sequence, word identification, jigsaw puzzle and memory task). According to the findings, dyslexic students have specific learning preferences and skills, as they tend to shift from declarative to procedural learning when experiencing stressful conditions. It was also found that there was a clear impairment for the declarative learning but no significant effect for the procedural learning. Regarding the results of the interview most of the 6 strengths were supported such as empathy, innovation and creativity.

Following the sequence of these two studies, I am going to investigate in depth the major causes and effects of stress on the performance of students and on teachers' role, respectively.

2. Literature Review

Stress is a physiological or psychological state which affects everybody at some point in their life. Pinel (2009) defines stress as "the physiological response to physical or psychological threat" (p.503). Stress is beneficial, to some extent, for the individual and improves the performance and quality of life since it contributes to individuals to experience challenges in their life (Tehrani & Ayling, 2008). However, when stress exceeds the limits it loses the beneficial qualities and is harmful for the individual (Cooper et al, 1988). Anxiety disorders are very common in childhood which leads to social and emotional problems, such as social inability and poor self-esteem and depression (Grills-Taquechel et al, 2012). According to Levine (2005) early experience can alter and affect people's behavioural and stress responses later in life. Depending on the nature of the stressor, if it is perceived as a challenge or a hindrance, students' performance is affected (Podsakof and Lepine, 2005).

It is argued that children with dyslexia show higher levels of anxiety and are more vulnerable to stress than normally achieving students. Casey et al's (1992) study found that dyslexic students have negative self-evaluation and are more depressed and sad compared to their peers as a consequence of reading disabilities. Bruck's (1989) study found that stress levels and anxiety reduce with age suggesting that as children leave school they are exempted from the pressure and heavy tasks of the educational system.

Reading and maths skills are undeniable useful in our everyday life, however there is evidence suggesting that the development of these skills is impaired because of high

levels of math and reading anxiety (Jansen et al, 2013). In particular, students with dyslexia are more overwhelmed by negative emotions and thoughts due to their literacy disability. Ashcraft and Krauze (2007) contend that maths anxiety is associated with low performance because anxious thoughts impair working memory and as a result it is difficult to use it to perform math tasks. Maths anxiety might also lead to low performance because sufferers try to finish the tasks as quick as possible, because they cannot deal with stressful situations (Ashcraft & Faust, 1994). Studies by Lee (2009) & Bandalos et al. (1995) suggest that interventions and success in maths result to higher levels of math and academic performance. Zbornik (2001) indicates that reading anxiety leads to physical and cognitive reactions, such as dread, phobia, low self-esteem and learned helplessness which impair students' performance in great extent. As a result, children with high reading anxiety have limited cognitive resources to cope with tasks of coding/decoding and understanding print. Grills – Taquechel et al (2002) report that literacy anxiety leads to poor academic achievement and can weaken the memory system and influence the information processing, specifically for maths and language tasks.

Teachers' role is very important in the classroom because they have to preserve the students' learning and psychological well – being. However, this role is so demanding that in most cases teachers suffer from stress, which can lead to mental and physical illness. It was suggested by Flook et al. (2013) that the major sources of stress include excessive workload, conflicts with managers and colleagues, poor students behaviour and lack of academic opportunities. It is very common that these factors push teachers to frequent burnout and turnover (Ingersoll, 2001).

Eide & Eide (2011) suggested the dyslexic advantage and argued for the gift of dyslexia. Dyslexia should not be seen as a disability only, but as a different way of thinking and performing. They argue that dyslexics develop coping strategies to stress and thus they are really good at problem solving and apply their dyslexic traits to cope with their difficulties. It is very comforting that dyslexics identify and work on their strengths and abilities.

3.   Proposed Methodology

The proposed study will investigate the causes and effects of stress in the classroom, regarding the performance of students and the successful teaching by the teachers. The aim of the study is to identify the major causes of stress which impair the success and high performance of teachers and students and develop classroom interventions which lead to a much improved learning. Therefore, we have firstly hypothesized that high levels of math and reading anxiety lead to low literacy skills and secondly that reducing levels of stress in classroom will improve students' performance and teachers' role.

To address these hypotheses we will conduct in-depth semi-structured interviews with primary school students and teachers and the data will be interpreted with the Interpretative Phenomenological Analysis (Smith 2008).

Students' interview will be based on the Math Anxiety Scale (AMAS) a psychometric property which measures math anxiety. There will be a sample of dyslexic and control students, aged 9-11 years old. The items applied in the interview will be related to the levels of anxiety students have when they think about an upcoming test sometime before, while watching the teacher work math problems on the blackboard, while taking an examination in math course or when being given a 'pop' quiz in math class. Equivalent questions will be addressed regarding the reading anxiety. A part of the interview will be related to potential strengths and observed weaknesses of dyslexia using the prototype Dyslexia School Strengths Finder by Agahi and Nicolson (2013). Topics to be explored will be big picture thinking, empathy, team work and creativity in school. This part of the interview aims to find what children enjoy, how they like to learn and how school corresponds with these preferences.

With regards to teachers, questions addressed in the interview will be related to their experience of stressful situations in classroom and the strategies or mechanisms used to tackle stress during teaching. The explored topics will focus on the excessive workload and working hours, on pressures from the authority or the school, on students' behaviour and indiscipline, and on the lack of career opportunities. The interviews will last approximately 30-45 minutes and will consist of 10 main and prompt questions for both teachers and students.

A set of four interventions effective with anxiety, suggested by Prof Rod Nicolson, will also be applied to students, according to the literature which suggests that with appropriate interventions the levels of stress and negative impacts are reduced (Brunye et al., 2013). Namely, there will be a brief focused-breathing relaxation exercise, which appears to be helpful to students to control their negative and pessimistic emotions prior to a math exam. An established method based on positive psychology, the three good things intervention, has proved effective in many situations, especially prior to numeracy and literacy exams. Third, exercise has proved beneficial in many domains, stimulating neuronal growth, with a study suggesting that there is a rebound effect after exercise that boosts declarative processing. Consequently, I will pilot a brief period of physical activity as a fourth intervention. A fifth intervention, for assessing comparative cost-effectiveness, involves 'class calming' in which disruptive individuals are provided with individually calming interventions. The interventions are designed to be immediately and easily usable and applicable by the teacher. They have been designed to facilitate different learning tasks in the classroom and therefore the time restriction has been

taken in mind so the interventions can easily be implemented during the one hour lesson. This study has been approved by the Department of Psychology Ethics Committee in accordance with the University of Sheffield's Research Ethics Approval Procedure.

4.   Conclusions

Although there have been many studies focusing on children with learning disabilities and behavioural problems, little research has investigated reading and math anxiety and academic achievements, specifically for children with learning difficulties. We hope therefore our research will be beneficial for the education system and shed light on the causes of negative emotions and feelings towards literacy skills and support students' and teachers experience and success, alleviated from stress. Teachers' role in preventing and helping students to overcome maths anxiety is very important, as they can use strategies and mechanisms to develop positive self-concepts and emotions to students (Fiore, 1999). Therefore, the scope of my study is to investigate the difficulties and stressful situations faced by teachers in the classroom and apply the interventions to students, which will be easily implemented in the classroom later by teachers as well. The findings seem to be promising as we hope to better assess, evaluate and help students who suffer from maths and reading anxiety. The perspective of "Positive Dyslexia" is also significant because it could have major implications on how the educational system deals with students with dyslexia and similar Learning Disabilities. It is significant to help students identify their strengths and potentials and cultivate their positive traits, as Humphrey (2002) suggests that self-perception affects how well we perform in school and in life, generally.

5.   References

Ashcraft, M. H., & Faust, M. W. (1994). Mathematics anxiety and mental arithmetic performance: An exploratory investigation. Cognition & Emotion, 8, 97-125

Bandalos, D. L., Yates, K., & Thorndike-Christ, T. (1995). Effects on math self – concept, perceived self – efficacy, and attributions for failure and success on test anxiety. Journal of Educational Psychology, 87, 611-623

Bruck, M. (1989). The adult outcomes of children with learning disabilities. Annals of Dyslexia, 39, 252-26

Brunye, T. T., Mahoney, C. R., Giles, G. E., Rapp, D. N., Taylor, H. A., & Kanarek, R. B. (2013). Learning and Individual Differences, 27, 1-7

Casey, R., Levy, S., Brown, K., & Brookes – Gunn, J. (1992). Impaired emotional health in children with mild reading disability. Developmental and Behavioural Pediatrics, 13, 256-260

Cooper, C. L., Cooper, R. D. & Eaker, L. H. (1988). Living with stress. Harmonsworth: Penguin

Eide, B, & Eide, F. (2011). The Dyslexic Advantage: Unlocking the hidden potential of the dyslexic brain. Hay House, Inc.

Fiore, G. (1999). Math – Abused Students: Are We Prepared To Teach Them? 92, 5, 403-406

Flook, L., Goldberg, S. B., Pinger, L., Bonus, K., & Davidson, R.J. (2013). Mindfulness for teachers: A pilot study to assess effects on stress, burnout and teaching efficacy. Mind Brain Educ., 7, 3

Grills – Taquechel, A. E., Fletcher, J. M., Vaughn, S. R., & Stuebing, K.K. (2012). Anxiety and Reading Difficulties in Early Elementary School: Evidence for Unidirectional – or Bi – Directional Relations? Child Psychiatry Hum, 43, 1, 35-47

Humphrey, N. (2002). Teacher and pupil ratings of self – esteem in developmental dyslexia. British Journal of Special Education, 29, 1

Ingersoll, R. M. (2001). Teacher Turnover and Teacher Shortages: An Organizational Analysis. American Educational Research Journal. 38, 3, 499-534

Jansen, B. R. J., Louwerse, J., Straatemeier, M., Van der Ven, S., Klinkenberg, S. & Van der Maas, H. (2013). The influence of experiencing success in math on math anxiety, perceived math competence, and math performance. Learning and Individual Differences, 24, 190-197

Lee, J. (2009). Universals and specifics on math self – concept, math self – efficacy, and math anxiety across 41 PISA 2003 participating countries. Learning and Individual Differences, 19, 355-365

Levine, S. (2005). Developmental determinants of sensitivity and resistance to stress. Psychoneuroendocrinology, 30, 10, 936-946

Nicolson, R. & Agahi, S. (2013). The Dyslexia Handbook. Positive Dyslexia: Working to our strengths? The British Dyslexia Association.

670

NUT Survey (2013)

Pinel, J. P. (2009). Biopsychology. Pearson Education.

Podsakof, N. P., LePine, J. A, & LePine, M. A. (2007). Differential challenge stressor – hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behaviour: a meta – analysis. Journal of Applied Psychology, 92, 2, 438

Smith, J. A. (2008). Qualitative Psychology. A Practical Guide to Research Methods. 2nd Ed.

Tehrani, N., Ayling, L. (2008). Work-related stress. CIPD Stress at work.

World Federation of Neurology. (1968). Report of Research Group on Dyslexia and World Illiteracy. Dallas: WFN

Zbornic, J. (2001). Reading anxiety manifests itself emotionally, intellectually. Today's School Psychologist.

Dyslexics and Their Entrepreneur Behaviors

Poliana Piacesi Sepulveda[1]

[1]Department of Psychology, University of Sheffield
ppsepulveda1@sheffield.ac.uk

Abstract. Many authorities have argued that the key focus of the entrepreneurial world in the new century will be attentive to specific traits, specifically those which could help to develop ones to become an entrepreneur being. Interestingly there appears to be a link between entrepreneurship and dyslexia, but this has not been fully explored using appropriate techniques. Academics have shown that dyslexic people show traits such as creativity, resilience, and delegation are really good at communication skills, which can turn them into persuasive beings. In this study I therefore undertook with a sample of 10 dyslexic entrepreneurs, a detailed qualitative study intended to identify the themes that clustered around entrepreneurial traits, dyslexic traits and those which were more generic traits for all entrepreneurs. 10 interviews were undertaken, 4 from the University of Sheffield, selected from a previous data-set, 6 from Brazil contacted through the Brazilian Dyslexia Association. 11 themes were identified which were communication skills, resilience, delegation/team work/shared tasks, vision, proactivity, risks with precaution, empathy, freedom, entrepreneur family, ask for help/modesty and control.

These categories of traits can be compared with the previous work of Logan (2009), one of the few publications in this field. Traits such as delegation and communication, which have already seen fit with the observations of Logan, however the ability in visual and spatial skills in areas like architecture, rich colour memory and the ability they have to use quick multi-sensorial conjunction are more related to the generic dyslexia traits (West, 1997 & Osmund, 1993).

Keywords: dyslexics, entrepreneur, traits, skills

1.   Introduction

Nowadays the terms entrepreneurship and entrepreneurs became topics very important for the world, such as political, cultural and economic (Clarke, 2007). Is a matter of fact that, according to Clarke (2007): ``entrepreneurship has been considered a principle means by which to enhance the creation of new jobs through the development of the small business sector (p. 6)´´. In such a big area of study,

which is entrepreneurship, this dissertation focus on the section of entrepreneurial behaviours. Along the dissertation, it will be discussed and researched linkages in between entrepreneur behaviours and dyslexic traits.

Several different academics have suggested different approaches for entrepreneurial traits. For example, Cromie (2000), explains that the trait approach suggests that people are already born with such traits and entrepreneurs are different from the rest of the population. One of the most recent studies in the field is from McClelland (1961), which explains how entrepreneurs act according to their need for achievement. There are also other approaches regarding this area, such as the Big Five (Costa and MaCrae, 1985).

This paper has the main aim to prove that they have entrepreneurial behaviours and they should take advantage of such trait, which can be used in their own benefit (opening their own companies) and/or also for companies that seek to employ people with entrepreneur spirit.

The latter will be the main focus of this study and it will be even more narrowed by attempting to apply principles of positive psychology in a workplace context relating to entrepreneur concerns, which will connect the main traits and behaviours of an entrepreneur according to Raunch & Freese (2007) and other authors, and also discuss it with some specific studies done by Logan (2001, 2009) in the field of dyslexics with dyslexia.

In order to address this topic first of all I will go through the literature on entrepreneurship followed by dyslexia, focusing primarily on the Julie Logan study (2009). However, in order to understand why dyslexia may be linked with entrepreneurialism, it is important to look at dyslexia and its main theories in more depth, which will be covered in the literature review.

Because of the relatively well-known difficulties associated with dyslexia, it can prove especially difficult for individuals in the workplace to find flourishing and motivating jobs. Often individuals with dyslexia feel excluded the society and may question what their strengths are.

In order to provide an adequate overview of Dyslexia I will first define it, then explain the medical model approach towards dyslexia, in the form of the main theories prevailing academic research approach to dyslexia followed by a recently more positive approach to Dyslexia termed Positive Dyslexia, which seeks to focus on the relative strengths of an individual with Dyslexia. According to an article written by Nicolson & Agahi (2013), Positive Dyslexia has three main focuses, which are ``positive experiences´´, such as previous happy experiences in life; ``Positive individual traits´´, which can include character and eventual interests by dyslexics and last, but not least ``Positive institutions´´ that are the community the dyslexic grew up, societies and family, for example. Nicolson & Agahi (2013) states that the main aim

of Positive Dyslexia is to ´´focus and strengthen these powerful ideas stimulating individuals with dyslexia to work to their strengths (p. 29).´´

One of the most popular themes brought out from Positive Dyslexia shows that there are clear strengths for Dyslexic adults in relation to entrepreneurial behaviours. Mainly supported by studies by Logan (2009) and also anecdotal evidence from amongst other research.

The work on entrepreneurship and Dyslexia has thus far been very useful however without sufficient depth analysis of the field and investigation into peoples subjective experiences. There have been questionnaires carried out, from which a number of interesting themes related to entrepreneurial behaviours have been selected and analysed. However something which needed further work was to understand people´s subjective experiences on how these entrepreneurial traits have manifested themselves into behaviours.

In order to guide this study my objectives are to identify in theory entrepreneur behaviours that dyslexic people can present, and to do so the study will start from an analysis of a secondary data previously collected by a PhD student in the Institute of Work Psychology. I be analysing the 'entrepreneurial trait' section of her secondary data set, which is based on a scale taken from Moriano et al. (2012), which lists 6 questions, each relating to a different trait (See Appendix). From this my sample 6 were selected for follow up interviews (See Methodology section). This was based on who had the highest scores.

This secondary data set was compiled from the development of a Dyslexia Works Strengths Finder, (DWSF) created with the intention of being able to assist individuals with Dyslexia get into careers more suited to their strengths. They contend that an individual with dyslexia (of whatever age) must have, as an integral part of their assessment, an analysis of their strengths as well as their weaknesses. The aim of the DWSF is to create a profile of skills which should prove the basis of a career development plan in which strengths are discussed, suitable career identified and then a development programme attain the planned outcomes.

In sequence, the qualitative research aims to deeper investigate which entrepreneur behaviours dyslexic adults strongly present with. The last objective is to identify the most common traits among the entrepreneur dyslexic sample and analyse it in the light of both entrepreneurship and dyslexia theories and previously studies.

Literature that will analyse and bring theories from general dyslexia studies to specific issues in dyslexia at the workplace. In the same section, theories concerning entrepreneurial traits and perspectives will also be discussed, in order to give a good background for the analysis, which will be the mixture between them.

Following the sequence of the study, the IPA (Interpretative Phenomenological Analysis) methodology will be presented and the specifications of data collection and

analysis as well. The next section will be focused in presenting the main findings from the qualitative research and personal thoughts and understandings about it. The 'findings' section will be followed by the 'discussion', which will cover the main findings in the light of the theory. This section will be divided into three parts, which will be a brief overview of main and most interesting findings, then limitations of the study and recommendations for future researches. Finally it will be given the conclusion, which will summarize all the main findings and most important discussed topics.

2. Literature Review

Nowadays people seek to be more entrepreneurial than ever. Aspirations towards this are mainly because of the freedom to work and the flexibility it gives. Others want to run away from a given and standardized structure, and there are even those who have the inner to achieve something unique and specific in their lives. For all these purposes, dyslexic individuals are highly likely to be entrepreneurs because of their early challenges when it comes to learning and standards.

Logan (2009) conducted a study with 139 adults, in which 102 were Entrepreneurs and the other 37 were Corporate Managers. Out of the 139 total samples, 39 adults were dyslexic and the 100 left were non-dyslexics. See table below:

Table 1: Sample composition from Julie Logan (2009) study

	Entrepreneur	Corporate Manager	Total
Non-dyslexic	66	34	100
Dyslexic	36	3	39
Total	102	37	139

Source: Logan (2009)

As it is clear to see, according to Logan's (2009) study, there is a higher incidence (92.3% probability) of being entrepreneur while being a dyslexic. Logan (2001) also argues that the probability of being entrepreneur is higher among dyslexic, than non-dyslexics. However, there is a small percentage of dyslexic being corporate managers.

Morgan & Klein (2000) and Morris (2002) confirm this result comparing the incidence of the American dyslexic sample (15%) with the reported incidence of being corporate manager, which is less than 1%. It is clear to see how bigger is the incidence of dyslexics being entrepreneurs (92.3%, Logan, 2009) comparing to corporate managers, which was found as less than 1% (Morgan and Klein, 2000; Morris 2002).

The view of Eide & Eide (2011) is really interesting when it comes to analyse dyslexia as being not just a learning disorder, but also an opportunity to recognize some common traits in dyslexics, which they can take as an advantage to develop better their activities. According to them, some of the important abilities dyslexic people present are: mechanical and relationships abilities, understanding abstract and metaphor situations.

All of these situations in where dyslexic people are privileged, also bring good connections for the entrepreneur world, with abilities such as creativity, proactivity personality, autonomy, innovativeness, among others (Raunch & Frese, 2007).

In the same direction of thought, Fitzgibbon and O'Connor (2002) states that dyslexic adults find coping strategies fundamental to overcome their weaknesses and these strategies make their entrepreneur side much stronger. The first coping strategy found by Logan (2009), Mazzarol (2003) and Timmons (1999) was delegation. According to them, this is a very important strategy, since dyslexics, in their majority, delegate tasks in which they feel week when performing, as an example, writing a memo. This strategy is interesting when analysed in two points of view: the first is the awareness that dyslexic people must have, which make them able to delegate the right tasks. And the second one is that the other ability they have to develop to delegate, which is relationship and trust, since in order to perform well when delegating, they have to have good relationship with other co-workers. Logan (2009) also states the importance to delegate while growing a business, because the right person would be fit in the right task, which could save money and time while executing tasks.

Logan (2009) also shows evidence of risk-taking being higher for dyslexic entrepreneur than for non-dyslexic ones. Rotter (1966) explains this strategy as being related to ``Internal locus of control´´, since people with it believe they can control outcomes by striving and persist more, in order to make it real. So, in their minds they have the thought of ``if I can control my outcome, I will take the risk´´, which definitely makes them people who put more efforts than non-dyslexics.

According to Raunch and Frese (2007), entrepreneur should have good generalised self-efficacy, since they have to be confidents in what they are doing, because in various situations they may be getting in uncertain situations and also taking some risks. People who have high self-efficacy are more likely to persevere and achieve their goals. In Logan's (2009) study, dyslexic entrepreneur were found with the same confidence level as non-dyslexic ones, which is a good indicator, since they do not feel shaken for having dyslexia, which can definitely help them to succeed faster.

Communication skills are usually very strong in dyslexic people. This is another coping strategy using by them, since it is a compensation mechanism for their lack of written skills (Nicolson and Fawcett, 1999). They also perceive themselves as stronger in this specific skill comparing to non-dyslexic entrepreneur. Communication is a crucial skill for being successful in the business world and this definitely helps them to inspire others and have/achieve a vision (Logan, 2009). Logan (2009) also

states that dyslexic entrepreneur take this skill as a very good advantage, since it facilitates networking, clarity when setting goals, establishing trust and also motivating those who work with them.

Dyslexic entrepreneur also have very unique traits and these ones definitely help them to succeed in the entrepreneurial world. According to Hales (1995) and Reid & Kirk (2001) these traits are: creativity, empathy and a very strong entrepreneurial orientation. These traits will be discussed among this dissertation.

Although dyslexic adults show so much interesting and valuable qualities to work in a company, many times they suffer the lack of support and also recognition from enterprises. In this scenario, dyslexics find really stressful to work for companies in this environment. This can definitely be found as one of the major reasons why dyslexics choose to start their own business, since they will be able to concentrate their qualities for what is really worth it (Hales, 1995: Reid \& Kirk, 2001).

As previously said, this environment makes dyslexics opt for entrepreneurship, since they can use their strengths, such as empathy and strong communication in their favour. In other situations they use their creativity by using coping strategies, which help them to overcome their weakness, for example by delegating a task they would struggle and lose too much time to complete (Logan, 2009).

3. Methodology

The research started by analysing a secondary data set from a PhD student in the Institute of Work Psychology. The aim of this was to determine the main characteristics and strengths of dyslexic adults and to help them get into careers suited for strengths. This data set was taken from the Dyslexia Work Strengths Finder (DWSF) which is a comprehensive tool aimed at investigating the various aspects of dyslexia strengths incorporating scales on Resilience, Self-Efficacy and Innovation to name a few. One of these sections of the DWSF was related to Entrepreneurial behaviours.   The scale where the data was retrieved was taken from Moriano et al. (2012). After analysing the secondary data, it was set the major characteristics that an entrepreneur should obtain and also the ones the study of Logan (2009) discovered. Those with the highest scores were contacted for follow up

The research was addressed in 10 interviews with a sample of dyslexic people (from Brazil and the United Kingdom) who considered themselves as entrepreneurs or having characteristics of one. The sample consisted in people with the mandatory characteristics previously stated and they were from contacts of my supervisors and also contacts from the Brazilian Dyslexic Association. The interviews were around 45 minutes long with 33 semi-structured questions composed of main questions and 'prompt' questions. The interviews were carried out either in person or via Skype. The approach used to analyse all my data in my research was the Interpretative Phenomenological Analysis (Smith, 2008).

The main purpose of this study was to analyse and confirm strengths and the relation that this could be used as an advantage for dyslexic who would like to be entrepreneurs. The focus of this study was in the Work Psychology area and discussing the entrepreneurial strengths at the workplace dyslexic people can present.

The following study had a sample of 10 participants, which were located in different countries: Brazil, England and The United States. Therefore, the consent form had to be signed and scanned. All participants agreed in taking part of the interview and agreed to be recorded. Participants from Brazil were contacted from the Brazilian Dyslexic Association. The participants from the United Kingdom were chose from the secondary data set which they highly ranked themselves as entrepreneurs in the questionnaire. The one participant from the US, was chosen by a personal contact, which had previously studied with this participant, when she was diagnosed with dyslexia.

4. Discussion

This research found eleven entrepreneur traits, which are present in the dyslexic's personality. From the summary of the findings in the following table, it is clear to perceive that there are some distinctive points about this dissertation results. Specific traits, which are modesty, family entrepreneurs and vision, were not broadly commented in the literature before, and some of them were not mentioned at all, for example modesty.

The following study had a sample of 10 participants, which were located in different countries: Brazil, England and The United States. Therefore, the consent form had to be signed and scanned. All participants

So far in this paper, we could confirm some of the research questions proposed in the beginning of this study. Until now, it is possible to say that dyslexics show entrepreneur behaviours; It was found new entrepreneur traits in dyslexic people, which were not previously shown in the literature and we could also outline some success in the interviewed sample about their success, since the majority of them have their own business, are planning to open it or already had.

The following table will give a brief overview of all the results previously presented from the research.

Table 2: Main entrepreneur traits found

Work-related Character Strengths	Cognitive Strengths	Inter-personal strengths
Proactivity (80%) Participants felt that they did not wait for an order to complete a given task. They also related proactivity with having a sense of freedom - having more autonomy helped them feel more proactive.	**Vision (60%)** The majority of the participants saw themselves as visionaries. They described themselves as looking for new opportunities and also long term perspectives, which, again, are crucial traits for successful entrepreneurs	**Delegation/Teamwork (70%)** Delegation, Team work and Shared Tasks were linked in all responses, leading to greater effectiveness and ease of operation
Risk-taking with precaution (90%) A key precaution was having the necessary 'knowledge' or expertise. Given this knowledge, they felt prepared to take risks, such as opening their own business, and also expanding it.		**Ask for Help (Modesty) (60%)** Some participants highlighted this as a crucial coping strategy for starting up their business plans and when they were uncertain
Resilience under pressure (75%) Participants felt that they worked well under pressure. With deadlines and being forced to be more focused they were able to work better		**Empathy (100%** Participants enjoyed helping and understanding others. Participants described how they felt a strong ability to recall emotions of others and by utilising this were able to understand the needs of those whom their business catered for.
		Communication Skills (90%) Participants considered they were very able to communicate and express their views and ideas . They also usedtheir communication skills as an advantage to compensate their writing skills, which are frequently quite poor.

Entrepreneurial Family (100%)
All participants had family members who owned their own businesses to some degree. Having these examples already in their lives gave them strong entrepreneurial tendencies themselves. In many cases because they had been working in those businesses themselves.
Control (40%)
Participants mentioned that their performance levels increased once they felt confident in having control over their work environments and the methods of how they worked.
Freedom (with focus and discipline) (100%)
All participants stated the importance of having the freedom to act flexibly and follow their instincts. However, the majority also emphasised the need for a routine and a schedule to keep them on track.

From all the main presented findings above, they are strongly related to entrepreneurial and dyslexics traits. In this section it will be discussed the importance of those results with the literature exposed above.

Logan (2009) states that dyslexics learn since an early age coping strategies, since they struggle so much until they reach their self-awareness of strengths and weaknesses. She also mentioned that some of these strategies very frequently are a good competitive advantage for the entrepreneurial world. Some of these traits, which became dyslexics' strategies from their self-awareness will now be discussed according to the research findings.

The first important trait found among in around 80% of the participants was the communication skills. This trait is fundamental for entrepreneurs, because whoever obtains such characteristic, will be most likely to have persuasion personality, which can increase the performance level of tasks in leading positions and in tasks associated with sells, for example (Siobhan, 2007). Siobhan (2007) also states that this relationship between persuasion and entrepreneurship is not only about communication, dialogue and a linguistic process, but rather a whole relational project, which: ``allows us to examine how entrepreneurs use a variety of means to create meanings and convince others of the validity of the venture (p. 91).´´

The same trait, according to Logan's (2009) study was found to be really strong in dyslexic, when compared to non-dyslexics entrepreneur. Logan's questionnaire asked the participants to rank themselves in communications skills. Non-dyslexics entrepreneurs, ranked themselves as average to good in this trait, while the majority of dyslexics considered themselves as 'very good'. Nicolson & Fawcett (1999) also argue that communication skills are part of dyslexics coping strategy to compensate their lack in the writing skills.

In this qualitative study, it was found that dyslexic entrepreneurs considered themselves as being really good when communicating and selling their ideas in the workplace. This finding was also connected with persuasion skills, according to these findings. Relating all the results, which concern communication skills, from the study, Logan (2009) confirms how important it is for a dyslexic to have such traits in the entrepreneurial world, by stating: ``Communication skills are an essential business tool. Those who can communicate well can inspire those around them to achieve a vision; they can network to build resources around an opportunity and motivate other to act.´´ Logan (2009) also suggests that such trait is a really good business advantage for entrepreneurs.

Communications skills are then, according with the exposed theory and findings, crucial for entrepreneur's success. Dyslexics who are entrepreneurs showed this trait

as being very consistent in their personality, which is a really good point for their development in the entrepreneurial world, because such trait can influence other characteristics such as network, persuasion skills, transforming the company's vision concrete and therefore the overall performance.

Another interesting result found in the taken research was the influence of the family into the dyslexics' entrepreneur career choice. All the participants in the sample stated they have a family member who is an entrepreneur. Moreover, in the majority of the cases they mentioned parents rather than relatives. Kets de Vries (1977; 1996) conducted a qualitative study with 40 participants and discovered that the family variable was a very relevant issue for the decision to become or not an entrepreneur. These results confirm Logan's (2009) finding in her study, which showed the positive influence from the dyslexics' family for their decision in opening their own business.

However, Logan (2009) found that the parents' influence is also related with the disorder as being hereditary, in the matter that their parents could also be dyslexics and therefore they could be the role model for their sons by running a successful family company. This finding does not corroborate with my findings, since just one participant stated that her sister was also dyslexic, which can just make us to conclude that entrepreneur family members are crucial for dyslexics' career choice, but dyslexic parents are not.

In the findings from the qualitative research from this dissertation, it was clear to see this trait very strong in the participants, which confirm both Logan's studies and findings and also make us conclude that dyslexics have this trait as an entrepreneur advantage, when it concerns to the speed in the company's growth.

We can now state that delegation is definitely a trait that dyslexic entrepreneurs have in advantage from the non-dyslexics entrepreneur, which is a really good indicator, since they have an easier time when it comes to work in teams. This delegation and team working facilities in their personality are also due to the fact they are very self-aware beings, according to the exposed above.

According to Raunch and Freese (2007), entrepreneurs have a really high internal locus control, which means that they think they means will predict their end, in other words, they feel they can control their outcomes at uncertain situations. In consequence from that, their risk taking status is high, since entrepreneurs believe they can prevent issues, by predicting possible issues.

In addition to that, Logan's (2009) study did find out a relationship between dyslexics and risk taking, which means this could also be another entrepreneur trait in the dyslexics' personality. In her study, non-dyslexics entrepreneurs ranked themselves as

'high level of risk taking', while dyslexics perceived themselves as 'very high on risk taking'. However, one of the participants who were in the dyslexic sample, was also diagnosed with ADHD, which could make the results unsure if such characteristic would be provoked by the dyslexia or the ADHD.

In the results found in this research, risk taking was another important finding from this paper research, which was found in the majority of the participants as well, which can definitely make us conclude that the relationship (dyslexia and risk taking) found by Logan (2009) was right, which can also not exclude the possibility of the other relationship (ADHD and risk taking). However, the answers were very straight and all of them said that they would all take risks, but with the some right precautions. Some of the precautions cited by them were: having the proper knowledge and a specific background, knowing the environment in where they were going, and being financially careful.

According to everything previously said, we can also link 'Control', which was an important finding in this paper' research, with the analysis about the internal locus control and risk taking. The participants stated that whenever they feel they have the control of the situation, they feel they perform their best and it is the most comfortable situation for them. This result just corroborates what Logan (2009) and Raunch and Freese (2007) said before about entrepreneur behaviours, since if they think they can control the outcomes, they will risk and feel more comfortable. Raunch and Freese (2007) also cited how control is related with the need for autonomy, because in entrepreneurs' mind, if they have this autonomy. They also stated: ``Need for control is also associated with entrepreneurs' avoidance of restrictive environments; they prefer to make decisions independent of supervisors, to set their own goals and develop their own plans and actions, and to control goal achievement themselves. People with high need of autonomy want to be in control (p. 359)´´.

Fitzgibbon and O'Connor (2002) also argue that dyslexics found really hard and stressful to work in companies with their own rules and standards. Therefore, they rather have situations where they can command their own days, determining their own day and controlling its outcomes. Taylor and Walter (2003) summarize it by saying that this is just another reason why dyslexics frequently choose to set up their own business.

Interestingly, 'freedom' was also another trait found in the research, which can confirm the statement above. It is absolutely clear to see how related control and need for autonomy (freedom) are and how it is a confirmation of an entrepreneur behaviour, which dyslexics strongly have, according to the results of this paper' research.

Connecting with 'risk taking' finding, the requirement for entrepreneurs to take risks is that they have to be in uncertain situations and to perform well in such situations; they have to be very resilient. According to Poon, Auinuddin and Junit (2006), stress tolerance is a key for a good performance when running its own business, because entrepreneurs have to take personal and financial risks constantly, and also overworking, which can be very stressful. The authors defined stress tolerance as the equilibrium state for successful entrepreneurs.

In the dyslexics' entrepreneurs, such trait was definitely seeing during the research and the majority of them both said they prefer to work under pressure and also perform better in this circumstances. They claim that stressful and pressure situations make them feel more focused and gives them more challenges.

As it was clear to see, risk taking and working under pressure were two results that came in different answers, however in the analysis they are vey related, because once dyslexics do not feel afraid to take diverse risks, they are consequently not scared to work in uncertain situations (according to their high internal locus control). Following the thought, once they have high internal locus control, they believe they can control their outcomes, so if they can do it, they are not afraid to face uncertain situations and also not work under pressure.

Masten (2001) and Ungar, Dummond and Mcdonald (2005) argues that a person who passed by a traumatic experience without giving up and also maintain his normal psyche, can be considered as resilient. Dyslexics, as we could demonstrate in along in this paper both in theory and in the research, suffered a lot with bullying and also with prejudice, and as we could perceive the majority of them were able to overcome the situation, in this way, it can be said dyslexics are resilient beings.

For dyslexics to have such trait, it is a really good indicator that they have another fundamental entrepreneur characteristic, which can give them advantages, such as working in extreme and different situations.

Johnson (2003) points out another personality trait' predictor for entrepreneurs, which is the empathy. He states that entrepreneurs very often are more inclined to be more person-oriented than task-oriented, which is a really good trait that entrepreneurs have as an advantage, since once you are leading a whole company, you have to make sure you are working in teams, because otherwise it can cost the price of a bad performance by the end of the month.

In the paper', empathy was definitely found as a trait present in dyslexics. And we can even explain it with a further thought. Since they have very strong communication skills, this can lead them to be more out spoken people, which can consequently have

an effect on how empathetic they are. Empathy can be then another quality of dyslexic entrepreneurs and also an advantage from their business competitiveness scenario.

According to Crant (1996), people who are highly ranked in proactivity want to be an influential person wherever they are. Frese and Fray (2001) also state that proactivity personality is something really intrinsic and individual for each person. According to them, proactivity is a fundamental trait, which entrepreneurs must have in their repertory, because such trait is related to influence environments when self-setting up their own business. Therefore, proactivity is then one of the most important traits that an entrepreneur should have, because entrepreneurs have to be the role model for their whole organization and if they do not give the first step, their employees will not.

Interestingly, in the research, this was a result, which was highly found among the participants. The majority of them considered themselves as being proactives, which is a really good signal and also confirms another crucial entrepreneur trait that dyslexics have.

After all those findings, a specific theme, which was ``risk with precautions´´ was brought into question and it also generated a hypotheses for future researches. It is really common for entrepreneurs to take risks, however, dyslexics showed their entrepreneur side (when assuming taking risks), but with precautions. The hypotheses in here would be if they do so, because of their past lived experiences, which created in their personality self-awareness to take precautions in their life.

A second hypotheses concern the ``modesty´´ theme, can be said as being a consequence of dyslexics' early life experiences when struggling with learning situations, specially reading. This specific situation could make them generate unconscious copying strategies, which could include asking help for people and in consequence become more modesty people.

The last hypotheses that came up was regarding the highly frequency of positive answers about having family member who are entrepreneurs. This finding could suggest that dyslexics can be highly influenced by relatives and parents when forming their personality traits when children, such as persuasion (communication), empathy, desire for freedom and proactivity.

To conclude this section, as we could perceive according to the theory and also the qualitative research, there are several entrepreneur traits that dyslexics obtain, and this can then suggest that dyslexia sample can be then another research sample for future studies concerning entrepreneurship topics. Going back to the research questions, it is perceived that all of them could be answered with presented findings. However, it was

also shown some new entrepreneurial traits in dyslexics, which could be deeper investigated in further researches and larger samples, in order to assure such traits in their personality.

5. Directions of further research

As we could perceive, there were some limitations among the present paper, which made some points not able to be more developed. Therefore, some of the recommendations for future researches and studies will be based on them.

During the research and the analysis of the specific literature review for dyslexia and entrepreneurship, we could not find any previous studies, which could support the findings from the trait modesty (ask for help). In the 'vision' trait, there was found a relationship with entrepreneur traits, however the same relation was not sustained by previous dyslexia studies.

Said that, the first recommendation from this study would be to further investigate those traits (modesty and vision), in both the entrepreneurial and dyslexia areas of study, in order to investigate the existence of such entrepreneur traits in dyslexic people.

It would be also interesting to do qualitative and quantitative researches in both dyslexics and non-dyslexics groups, with the focus to study whether or not dyslexics would be better entrepreneurs, since they show more self-awareness and know better their limits, because of lived experiences.

Another issue, which was stated as a limitation was the few sample I was able to work with, which influence the validity of some findings. In order to develop this problem, it is suggested that in future studies, academics take part in a quantitative research with a sample that can that will be enough to give validity from new possible entrepreneur traits found in dyslexics, and also to confirm those, which were already mentioned before from other academics.

6. Conclusions

According to everything previously mentioned, there is an impressive list of entrepreneur traits, which dyslexics presented during the qualitative research, and not only that but the majority of them were already predicted by Logan (2009), Nicolson & Fawcett (1999) and others. Such initial findings answer and confirm dissertation' general objectives and research questions proposed in the beginning of the study.

The most important traits found by the research were the influence to choose the entrepreneurial career by dyslexics' family; freedom, since they do not like any ready

and standards structures in their workplace; Empathy was another trait very much mentioned by the sample, and it has the connection with how person-oriented they are at the workplace and how this is important when an entrepreneur aims to form teamworking at their companies; Risk with precautions was another trait found due to their high internal locus control, already discussed above. Proactivity was identified by the participants and theory as well as being entrepreneur behaviours presented in dyslexics, which means that they are beings who are self-starts and seek for opportunities.

Again confirming the thought of Masten (2001) and Ungar, Dumond & Mcdonald (2005), dyslexics were found in the dissertation's results as resilient beings, because of possible traumas they suffered in their early age and turned it into their strengths, which definitely helped them when being entrepreneur, since they stated how comfortable they feel when working under pressure and also in general uncertain situations. Dyslexics were also found to be strong in delegation, which highly contributed for team working and shared tasks. These traits were found both in literature and also in the qualitative research results. Such traits are crucial for entrepreneurs, since they have to be self-aware of weaknesses and strengths and be able to delegate the right tasks to the right people to execute them.

Vision was a trait very much mentioned by academics, such as Raunch & Freese (2007), which they stated the importance of having it for entrepreneurs and how it is fundamental for them, since it is the first step to set goals, see new opportunities and go for challenges. The same trait was highly shown among the dyslexic sample. However, none of the literature that concerned dyslexics issues talked about it.

The traits ask for help and modesty was also found among the dyslexic sample used in the research and although we can see a relationship in between the dyslexic part, which is that since early age they become self-aware of weaknesses and have to develop copying strategies, such as this one; and also in the entrepreneurial side, which is very related to growth and how important it is to ask for help to partners, co-workers, experts by using its network. This trait was not discussed in previous studies concerning neither entrepreneurship nor dyslexia.

And the last, but not least was the communication skills trait, which was definitely predicted by authors in both areas, dyslexia and entrepreneurship. Dyslexics have such trait, since they find it a way to compensate their writing skills. Raunch and Freese (2007) also argue that being good at communication is crucial for an entrepreneur, since such ability is related to persuasion and for entrepreneurs, it is another trait pretty much needed.

7. References

Agahi, S. & Nicolson, R. (2013). The Dyslexia Handbook. Positive Dyslexia: Working to our strengths?. The British Dyslexia Association

Baum, J. R., & Locke, E. A. (2004). The relation of entrepreneurial traits, skill, and motivation to subsequent venture growth. Journal of Applied Psychology, 89(4), 587-598.

Clarke, J. (2007). Seeing entrepreneurship: Visual Ethnographies of embodied entrepreneurs. Leed's University Business School.

Costa, P. T. & McCrae, r. r. (1985). The NEO Personality Inventory Manual. Odessa, FL: Psychological Assessment Resources.

Crant, J. M. (1996). The proactive personality scale as a predictor of entrepreneurial intentions. Journal of Small Business Management, 34(3), 42-50.

Cromie, S. (2000). Assessing entrepreneurial inclinations: Some approaches and empirical evidence. European Journal of Work and Organisational Psychology. 9(1). 7-30

Cromie, S. (2000). Assessing entrepreneurial intentions: Some approaches and impirical evidence. European Journal of Work and Organizational Psychology, 9(1), 7-30.

Eden, D. (1988). Pygmalion, goal setting, and expectancy: Compatible ways to raise productivity. Academy of Management Review, 13(4), 639-652.

Eide, L. & Eide, F. (2011). The dyslexic advantage: unlocking the hidden potential of the dyslexic brain. Hay House.

Fawcett, A. J., & Nicolson, R. I. (1992). Automatisation deficits in balance for dyslexic children. Perceptual and Motor Skills, 75, 507-529

Fawcett, A. J., & Nicolson, R. I. (1999). Performance of Dyslexia children on cerebellar and cognitive tests. Journal of Motor Behaviour, 31, 68-78

Fitzgibbon, G, & O' Connor, B. (2002). Adult dyslexia: A guide for the workplace. London: Wiley

Gartner, W. B (1985). A conceptual framework for describing the phenomenon of the new venture creation. Academy of Management Review, 10(4), 696-706

Gartner, W. B (1989). ``Who is an entrepreneur?´´ is the wrong question. Entrepreneurship Theory and Practice, 12(2), 47-68.

Hales, G. (1995). ``Stress factors in the workplace´´. In Miles, T.R. \& Varma, V. (eds) (1995) Dyslexia & Stress. London: Whurr. pp. 74 – 82.

Hoffman, F.J., Sheldon, K.L., Minskoff, E.H., Sautter, S.W., Steidle, E.F., Baker, D.P., Bailey, M.B and Echols, L.D. (1987). Needs of learning disabled adults. Journal of Learning Disabilities, 20 (1), 43-52.

Kets de Vries, M. F. R. (1977). The entrepreneurial personality: a person at a crossroads. Journal of Management Studies, 14(1) 34-57. 34-57.

Kets de Vries, M. F. R. (1996). The anatomy of the entrepreneur: clinical observations. Human Relations. 49(7), the 853-883

Knight, F. H. (1921). Risk, uncertainty, and profit. New York: Kelly \& Millman.

Logan, J. (2001). ``Entrepreneurial success: A study of the incidence of dyslexia in the entrepreneurial population and the influence of dyslexia on success´´. PhD thesis, Bristol University

Logan, J. (2009). Dyslexic Entrepreneurs: The incidence, Their Coping Strategies and Their Business Skills. Wiley InterScience.

Masten, A. S. (2001). Ordinary magic: Resilience processes in development. American Psychologist 56(3) 227-38.

Mazzarol, T. (2003). A model of small business HR management. International Journal of Entrepreneurship Behaviour \& Research, 9(1), 27-49.

McClelleland, D. C. (1961).The Achieving Society. Princeton, NJ: Van Nostrand.

Morgan, E., & Klein, C. (2000). The dyslexic adult in a non-dyslexic world. London: Whurr.

Moriano et.al. (2012). A Cross-Cultural Approach to Understanding Entrepreneurial Intention. Journal of Career Development.

Morris, B. (2002). Overcoming dyslexia. Fortune Magazine, May 13.

Nicolson, R. I., & Fawcett, A. J. (1990). Automaticity: A new framework for dyslexia research? Cognition, 35, 159-182

Nicolson, R. I., & Fawcett, A. J. (1999). Developmental dyslexia: The role of the cerebellum. Dyslexia: An International Journal of Research and Practice, 5(3), 155-177.

Osmund, J. (1993) The Reality of Dyslexia. London: Cassell.

Plata, M. and Bone, J. (1989) Perceived importance of occupations by adolescents with and without learning disabilities.

Poon, J. M. L., Ainuddin, R. A., \& Junit, S. H. (2006). Effects of self-concept traits and entrepreneurial orientation of firm performance. International Small Business Journal, 24(1), 61-82.

Raunch, A., Frese, M (2007). Let's put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners' traits, business creation, and success. European Journal of Work and Organizational Psychology, 16(4), 353-385.

Reid, G, Kirk, J (2001). Dyslexia in Adults, Education and employment. Wiley

Rotter, J. B. (1966). Generalised expectancies for internal versus external control of reinforcement. Psychological Monopraphs, 80(1, Whole No. 609).

Rotter, J. B. (1966). Generalised expectancies for internal versus external control of reinforcement. Psychological Mongraphs, 80(1), 1-2.

Sarasvathy, S. D. (2001). Causation and effectuation: Toward a theorethical shift from economicinevitability to entrepreneurial contingency. Academy of Management Review, 26, 243-263.

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. American Psychologist, 55, 5–14.

Shane, S. (2003). A general theory of entrepreurship: The individual-opportunity nxus. Northhampton, MA: Edward Elgar Press.

Shane, S., & Venkataraman, S. (2000). The promise of entrepreneurship as a field research. Academy of Management Journal, 25(1), 217-226.

Siobhan, J. C. (2007) Seeing entrepreneurship : visual ethnographies of embodied entrepreneurs. PhD thesis, University of Leeds.

Smith, J. A (2008) Qualitative Psychology. A Practical Guide to Research Methods. 2nd Ed.

Smith, J.A., Flowers, P., & Larkin, M. (2009). Interpretative Phenomenological Analysis: Theory Method and Research. London: Sage.

Stein, J. (2001). The magnocellular theory of developmental dyslexia. Dyslexia, 7, 12-36

Stewart, W. H. & Roth, P. L. (2001). Risk propensity differences between entrepreneurs and managers: A meta-analytic review. Journal of Applied Psychology, 86(1), 145-153.

Taylor, K. & Walter, J. (2003). Occupation choices of adults with and without dyslexia. Dyslexia: An International Journal of Research and Practice, 9(3), 177-185.

Timmons, J. (1999). New venture creation (5th ed.). Singapore: Mc-Graw-Hill..

Ungar M., Dumond C. and Mcdonald W (2005). Risk, Resilience and Outdoor Programmes for a risk children'. Journal of Social Work. 5(3) pp. 319-338.

Vellutino, F. R., J. M. Fletcher, et al. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? Journal of Child Psychology and Psychiatry 45: 2-40.

West, T. G., (1997.) In the Mind's Eye: Visual Thinkers, Gifted People with Dyslexia and Other Learning Difficulties, Computer Images, and the Ironies of Creativity. Amherst, N.Y.: Prometheus Books.

West., T.G., (2009). In the Mind's Eye: Creative Visual Thinkers, Gifted Dyslexics, and the Rise of Visual Technologies. Amherst, MA: Prometheus Books

World Federation of Neurology. (1968). Report of research group on dyslexia and world illiteracy. Dallas: WFN.

Differentiated Instruction in the Context of the Greek State Primary School: An Exploratory Study of Teachers' Attitudes and Classroom Implementation

Anastasia Mavroudi[1]

[1]PhD Candidate, Department of Theoretical and Applied Linguistics, School of English,

Aristotle University of Thessaloniki
mavana500@gmail.com

Abstract. The present paper investigates the perceptions of differentiated instruction of Greek state school teachers of English. It also sets out to discover whether they use differentiated strategies in their classes and which techniques they regard as more appropriate for their teaching context. The effect of factors such as the teachers' age and academic training on their attitudes towards differentiated instruction and on their use of differentiated strategies is explored as well. For this purpose, a self-report questionnaire was completed by 149 teachers of English working in Greek state primary schools. The findings indicate that the respondents' overall attitude towards differentiated instruction is positive. However, although several features of differentiated instruction have been adopted, its full implementation is hindered by misconceptions on the part of the teachers and practical considerations, such as lack of training, time and resources. The findings presented in this paper are derived from on-going research for a doctoral thesis on Greek state school teachers' perceptions and use of differentiated strategies.

Keywords: Differentiated instruction, differentiated strategies, English language teaching in Greece

1   Introduction

In recent years, significant changes in classroom populations have been noted in Greece, as well as worldwide. Within the same classroom, there may be variance in the students' cognitive, affective and physical development, their level of proficiency, their socio-economic background and in many other aspects. Such diversity in classroom populations creates new challenges and opportunities for today's educators, since learners who differ so greatly from one another cannot be expected to assimilate the same subject matter, in the same way and at the same pace (Tomlinson, 1999). The idea of Differentiated Instruction has been proposed in response to administrators', teachers', students' and parents' concerns with classroom heterogeneity. The present paper examines certain basic characteristics of Differentiated Instruction and a number of differentiated teaching strategies. More importantly, it seeks to discover which of these strategies are currently used by teachers of English working in Greek state

primary schools, according to their own statements in a self-report questionnaire, as well as the teachers' overall attitude towards Differentiated Instruction.

2   Literature Review

For the purposes of the present paper Benjamin's definition of Differentiated Instruction (DI) is adopted i.e. DI is described as "a broad term that refers to a variety of classroom practices that allow for differences in students' learning styles, interests, prior knowledge, socialization needs, and comfort zones" (2003, p. 1). If implemented properly, DI enables educators to maximize each learner's growth and individual success, by discovering what he or she already knows (i.e. their point of entry in the learning process, or what is often called "readiness") and building on from there, along a continuum (Theisen, 2003). DI would thus promote academic excellence and respect of learner heterogeneity for the students.

Bearing in mind their students' readiness or level of attainment, their learning styles and interests, teachers can design lessons in a way that fits the needs of individual learners, instead of "teaching to the middle". Curricular elements that can be differentiated include the *content, process* and *product* of the lesson (Theisen, 2003). These terms will be briefly explained due to their relevance to the findings presented in this paper.

Content refers to the input of the lesson e.g. the ideas, facts or information communicated to the learners. Providing struggling students with additional examples and more opportunities for practice, while more proficient students work on more complex tasks, is an instance of content differentiation. Similarly, students may be given multiple versions of the same reading text, varied in terms of difficulty, or different reading texts, some of which may be simpler and others more challenging.

Process differentiation takes place when different ways of assimilating the input of the lesson are suggested to the learners. Students may work individually, in pairs or in groups in order to learn more effectively. Flexible grouping should be applied as well, as it keeps students constantly challenged (Vastaki, 2010) and provides them with the opportunity to assume different roles within groups and to practice different skills. The pace of the lesson may be varied to better suit individual learners' level of attainment and keep them motivated (Vastaki, 2010). Students should be provided with many options on how to learn and with a variety of resources. In this way, student autonomy develops and they assume more responsibility for their own learning. This has been said to maximize learning (Theisen, 2003).

Finally, differentiating the product of a lesson implies that the output of the lesson will not be the same for all students. Students may be provided with a list of options for homework, so as to choose the one that best suits their interests or level of ability. Varied tests may be considered as another instance of product differentiation (Theisen, 2003).

Regarding the effectiveness of DI, studies on the impact of DI on student learning have shown it to have a beneficial effect on both gifted students (Tieso, 2005) and on children with learning disabilities (McQuarrie, McRae and Stack-Cutler, 2008). The same applies to various age groups e.g. primary school students (Bedee, 2010) or middle and high school learners (Rasmussen, 2006). These studies provide the

empirical evidence to back up the claim that DI is an effective approach that embraces learner heterogeneity and helps all students reach their full potential.

With regard to the effect of teachers' personal characteristics on their attitudes towards DI, research suggests that age and academic training may play a significant role. Several researchers state that older teachers are less positively predisposed towards new ideas and practices (Avidov-Ungar and Eshet-Alkakay, 2011, Kusano et al., 2013). Moreover, claims have been made that teachers with a more extensive formal education tend to be more positively predisposed towards innovative ideas. For example, Burns (2005) discovered that holders of postgraduate degrees were less resistant to differentiated instruction than participants with fewer formal qualifications.

As far as Greece is concerned, DI has been introduced relatively recently in the Greek educational system and can be said to constitute an innovative idea. The teaching of English is carried out according to the guidelines of the Cross-Thematic Curriculum for the teaching of English in Greek state schools (Presidential Decree Φ.Ε.Κ. 303/10-03-03). Published in 2003, the curriculum placed great emphasis on the fact that each learner is different and that individual learners' differences must be taken into account, so that all students are allowed to make the most of the educational process.

A new curriculum for the instruction of all foreign languages taught in state schools across the country was published in September 2011, the Integrated Foreign Languages Curriculum (IFLC). It is implemented on an experimental basis on a selected number of schools, but is expected to become the official curriculum later on. According to the IFLC, DI is regarded as essential in dealing with classroom heterogeneity (*Foreign Languages at School: Guide for the Foreign Languages Teacher*, 2011). Since it is the responsibility of the state school teachers to transform the mandates of the two curricula into effective classroom practices, it is important to examine whether educators have sufficient knowledge and understanding of DI principles and whether they have adopted differentiated strategies in their teaching.

3   Methodology

Purpose of the Study and Research Questions

The aim of this study is to investigate the extent to which and the ways in which DI has been incorporated in EFL teachers' classroom practices in Greek state primary schools. For this purpose the following research questions need to be answered:

• Do EFL teachers in Greek state primary schools have a clear understanding of the rationale and principles of DI?
• Has DI influenced primary EFL teachers' classroom practices? Do teachers use differentiated techniques in the classroom and which ones do they find more appropriate?
• Do factors such as the teachers' age and educational background affect their use of DI techniques?

Participants and Context of the Study

The participants are 149 English teachers who work in Greek state primary schools in the region of central Macedonia, i.e. in northern Greece. Central Macedonia was selected due to its being a densely populated region, which includes both urban and rural areas. As a result of the population size, geographical features and diversity of the region, it can be said that the conditions found in this region closely resemble those in the rest of the country. In this way, valid conclusions can be drawn for Greek state primary schools in general.

Moreover, English language teaching in state schools of the region is typical of EFL teaching throughout Greece, given that decisions regarding the curriculum, materials, hours of instruction, etc. are made at Ministry level and apply to all schools in the country. Teachers in central Macedonia are similar to their colleagues working in other parts of Greece in terms of important characteristics such as age, gender, ethnicity, educational background, formal qualifications and socioeconomic status. Thus, any conclusions drawn from the study may be expected to be generalizable to the entire population of primary EFL teachers in Greece.

Instrumentation

For the purposes of the study in question, a questionnaire was designed, so as to guarantee that a sufficient and easily analyzable amount of data is collected (Dörnyei, 2003). The questionnaire included various parts, but the first and second part will constitute the focus of the present paper.

The first part of the questionnaire was largely designed as a Likert scale aiming to elicit the compatibility between respondents' views and DI tenets and the degree of the respondents' understanding of DI underlying principles through the expression of agreement or disagreement to a number of statements.

In the second part of the questionnaire, respondents were requested to reflect on their teaching and to state the frequency of use of differentiated techniques in their own classrooms. A five-point scale was employed (5: always, 4: very often, 3: quite often, 2: seldom, 1: never). Personal information on the respondents' age, gender, educational background and teaching experience was also drawn.

Data Collection and Analysis

The questionnaires were collected over a 7-month period. Bearing in mind that 149 teachers took part in the study, participation amounted to 23.5% of the target population (637 teachers). Moreover, measures were taken in order to ensure that all areas in the region were fairly represented. As a result, it could be said that the sampling procedure used is what Dörnyei (2003, p. 73) terms *area/ cluster/ stratified random sampling*; the population was divided into groups according to the area they work in and a random sample from each group was subsequently selected. The data were analyzed using SPSS.

4 Results

Teachers' Attitudes towards Differentiated Instruction

The first section of the questionnaire dealt with teachers' attitudes towards DI. Three areas of interest were identified, namely, differentiating the pace of the lesson, differentiating the activities the students work on and developing students' autonomy as learners. Teachers' beliefs and views on these topics were then explored.

Differentiating the Pace of the Lesson

To begin with, respondents are well aware of classroom heterogeneity. The number of respondents who disagree with the idea of all students learning the same things, at the same time and in the same way is overwhelming (98%). However, differentiating the pace of the lesson to accommodate different learners is not unanimously accepted. 46.2% either strongly or mildly disagree with the idea that all students should go through the materials to be learned at the same rate, but 53.8% of respondents agree with keeping the pace stable for all learners. Acceleration, a strategy whereby students who assimilate the materials at a faster rate are allowed to move on to new information, is viewed in a negative light by 71.4% of respondents. However, respondents are not as unfavourably predisposed towards the idea of having struggling students move at a slower pace. 24.1% foster negative attitudes towards this strategy, but most respondents (76.2%) agree with slowing the rhythm down for struggling students.

Finally, as far as whole classroom instruction is concerned, that is, the mode of instruction whereby the teacher works with all the students at the same time, 38.3% of respondents express varying degrees of disagreement, whereas 61.7% agree or strongly agree with whole classroom instruction. The findings described above are presented in the following table (table 1).

Table 1. Pace of the lesson

PACE OF LESSON	Strongly Disagree	Disagree	Don't know	Agree	Strongly Agree
Students can learn the same things, at the same time, in the same way	71.8% 107	26.2% 39	0% 0	1.3% 2	0.7% 1
Same pace for all students	5.4% 8	40.8% 60	1.3% 2	49.7% 73	4.1% 6
Acceleration of pace	10.7% 15	60.7% 85	4.7% 7	27.9% 39	0.7% 1
Whole class instruction	1.4% 2	36.9% 52	2.7% 4	54.6% 77	7.1% 10
Slower pace for struggling students	0.7% 1	23.1% 33	2% 3	62.2% 89	14% 20

Differentiation of Activities

With regard to the differentiation of learning materials, the picture is clearer. 72.1% of the respondents foster positive attitudes towards their students being given different tasks to do depending on the learners' level of ability. An additional 15% strongly agree with differentiating the materials according to their students' level of attainment. Moreover, respondents were requested to express their views on the idea of all students working on the same activities, while the teacher provided additional support to struggling learners. 61.5% of respondents agree and 31.1% strongly agree with it. The number of respondents who disagree with this statement is considerably lower (7.5%). To sum up, differentiation of the contents of the lesson is more easily accepted by respondents compared with differentiation of the pace of the lesson. These findings are presented in table form below (table 2).

Table 2. Activities

ACTIVITIES	Strongly Disagree	Disagree	Don't know	Agree	Strongly Agree
Different activities based on students' level of ability	1.4% 2	11.6% 17	0% 0	72.1% 106	15% 22
Same activities, additional support for struggling students	0.7% 1	6.8% 10	0.7% 1	61.5% 91	31.1% 46

Developing Students' Autonomy as Learners

With regard to the issue of learners' autonomy, respondents were asked whether the teacher could assign tasks to the students to work on, while s/he could act as monitor and only help when necessary. Thus, a portion of the teacher's authority and control would be passed on to the students. Respondents expressed equal amounts of agreement (48.6%) and strong agreement (49.4%) to this idea. Only 2% of respondents disagreed with this idea. However, this view of the role of the teacher is in conflict with the statement on whole classroom instruction, whereby the implied teacher role is that of an omniscient authority. These contradictory findings will be discussed in the subsequent section (section 5).

Differentiated Strategies

The second section of the questionnaire was designed with the aim of eliciting information on the respondents' teaching strategies. The respondents were requested to state how frequently they make use of various differentiated strategies, e.g. the provision of differentiated reading texts, in their own classrooms. Four areas of

interest were focused on, namely, group work, pair work, written tests and reading materials.

Group Work

As far as group work is concerned, there are two options. Students could be grouped according to their level of attainment or according to shared interests. Respondents seem to prefer grouping their students according to the latter rather than creating groups based on the learners' level of attainment. Only 1.4% of the teachers claim to always group students according to level of ability. The majority of respondents opt for this technique either rarely (36.8%) or never (29.2%). On the contrary, 41% of respondents state that they usually group students based on their interests and 34% often do so.

Table 3. Group work

GROUP WORK	Always	Usually	Often	Rarely	Never	Total
Students grouped according to level of ability	1.4% 2	10.4% 15	22.2% 32	36.8% 53	29.2% 42	100% 144
Students grouped according to interests	14.6% 21	41% 59	34% 49	6.9% 10	3.5% 5	100% 144

Pair Work

Respondents were asked to reflect on pair work in the classroom and on how the learners' level of attainment affects the way they pair their students. 47.1% state that they often pair students with the same level of ability, whereas 27.3% claim to rarely do so. Pairing students with a different level of ability is clearly very popular among Greek state school teachers; the majority (48.9%) of respondents claim to use this strategy in either every pair work activity or in most pair work activities. 38.8% also state that they often pair their students in this way. Nevertheless, although teachers generally view group work and pair work favorably, they often ask their students to work individually (54.3%). Moreover, 34.9% of the respondents claim that their students usually work on their own.

Table 4. Pair work

PAIR WORK	Always	Usually	Often	Rarely	Never	Total
Same level of ability	1.7% 2	19% 23	47.1% 57	27.3% 33	5% 6	100% 121
Different level of ability	8.6% 12	40.3% 56	38.8% 54	10.1% 14	2.2% 3	100% 139

Individual work	2.3% 3	34.9% 45	54.3% 70	7.8% 10	0.8% 1	100% 129

Written Tests

The topic of students' evaluation via written tests was also examined. Respondents were asked to define how frequently they produce one version of a written test to be administered to the whole class, as opposed to producing multiple versions of a test, with exercises of graded difficulty for different groups of students. The majority (61.6%) claim that they always use one test for all learners, whereas the percentages of those who rarely or never produce a single test for all learners are very small (2.7% and 1.4% respectively).

Table 5. Written tests

WRITTEN TESTS	Always	Usually	Often	Rarely	Never	Total
Same test for all students	61.6% 90	23.3% 34	11% 16	2.7% 4	1.4% 2	100% 146

Differentiation of Reading Materials

Respondents were also requested to comment on their choice of reading materials for their students. The coursebooks invariably used in the state school classrooms do not contain multiple reading texts for students with different levels of attainment, interests or learning profiles. It is up to the teachers to produce graded versions of the same text or to provide students with many texts depending on the topics that appeal to them. However, this task appears daunting to most teachers. 31.4% state that they always use the same text with all students in a particular class, 41.6% claim that they usually work with one reading text and 21.9% often do so. Only one respondent (0.7%) states that his/her students never work on a single text.

Table 6. Criteria for the differentiation of reading materials

DIFFERENTIATION OF READING MATERIALS	Always	Usually	Often	Rarely	Never	Total
Same reading texts	31.4% 43	41.6% 57	21.9% 30	4.4% 6	0.7% 1	100% 137

The Impact of Teachers' Age and Education on their Attitudes Towards Aspects of DI

The impact of teachers' age and academic training on their attitudes towards DI and on their adoption of differentiated strategies is examined through the inferential statistics presented in tables 7 and 8. As can be seen in the former table, with regard to an aspect of flexible grouping whereby some students may work in groups while others could work on their own, the vast majority of participants holding a basic university degree disagree with this idea (72.2%). Only 27.8% of them are positively predisposed towards this type of flexible grouping. On the other hand, approximately half (47.5%) of the participants who have completed a Master's or PhD program agree with having various types of student groupings working simultaneously in the classroom and the other half (52.5%) disagree with this practice. Thus, it can be said that holders of a postgraduate degree are more open to using this type of flexible grouping with their classes than participants with a basic university degree.

Table 7. Participants' education and types of flexible grouping

	SOME STUDENTS WORK IN GROUPS, OTHERS ON THEIR OWN		Total
	A/SA	D/SD	
BASIC UNIVERSITY DEGREE	27.8% 30	72.2% 78	100% 108
MASTER'S/ PhD	47.5% 19	52.5% 21	100% 40
Total	33.1% 49	66.9% 99	100% 148

$\chi^2 (1) = 5.127, \ p = 0.024$

As far as the impact of teachers' age on their choice of classroom techniques is concerned, two classroom practices are examined here, namely, producing a single version of a written test for all students and differentiating reading materials based on the students' level of attainment. With regard to producing only one version of a written test for all students in a classroom, it can be observed that younger participants (of 30 years of age or less) disagree with this practice to a greater extent than their colleagues aged 41-50. The former age group also claims to provide the same test for all students less frequently than participants aged 51 or more do.

With regard to differentiating reading materials according to the students' level of ability, statistically significant findings in table 8 reveal that participants of 30 years of age or younger report that they practice differentiation of reading materials more often than do teachers aged 51 or more. On the whole, it can be said that the younger

teachers are, the more frequently they claim to use these strategies, which are consonant with DI tenets.

Table 8. The impact of participants' age on classroom practices

	AGE GROUP	N	M	SD	ANOVA	Tukey HSD
SAME TEST FOR ALL STUDENTS	- 30	10	2.40 *	0.96	F(3.144) = 5.35 p = 0.002	- 30 vs 41-50 p = 0.023 - 30 vs 51 + p = 0.001
	31-40	43	1.67	0.97		
	41-50	69	1.57 *	0.88		
	51+	23	1.13 *	0.34		
	Total	145	1.59	0.89		
DIFFERENTIATION OF READING MATERIALS ACCORDING TO STUDENTS' LEVEL OF ABILITY	- 30	10	2.70 *	1.05	F(3.140) = 2.65 p = 0.052	- 30 vs 51+ p = 0.032
	31-40	40	3.40	1.03		
	41-50	69	3.38	1.17		
	51+	22	3.86 *	0.99		
	Total	141	3.41	1.12		

* denotes statistical differences between age groups

5   Discussion

DI Underlying Principles

The findings provide evidence that respondents in this study generally foster a positive attitude towards basic DI tenets. They acknowledge that the learners' unique characteristics need to be taken into account in order for instruction to be effective. Respondents also appear to be particularly sensitive to the fact that there may be students with different levels of attainment within the same classroom. Thus, they express their disagreement with having all students work on the same materials, stating that either the activities used should be differentiated to match the students' different levels of ability or more help should be offered to struggling students.

   Differentiation of pace is also seen as a helpful technique by many respondents, but the majority believe the strategy of acceleration, which is one of the forms that differentiation of pace can take, to be a possible source of problems. It may be the case that the preparation of special materials required for the highly proficient students is viewed as extra workload. Teachers may also think that having some students go through the materials at a faster rate can dishearten struggling students that feel they cannot keep up with their classmates. It is also possible that teachers are worried that acceleration may lead to their losing their authority in the classroom, since they would

no longer control the pace of work for all students, and it could thus result in discipline problems. This is also supported by the respondents' answers with regard to whole class instruction, which is undeniably popular among teachers and enables them to better control their classes.

On the other hand, allowing struggling students to work at a slower pace seems to be well-received, as three out of four respondents agree with the use of this technique. As a result, it can be said that differentiating the pace of the lesson causes mixed feelings to Greek state primary school teachers of English, depending on the form this strategy takes.

With regard to the respondents' attitudes towards the role of the teacher as monitor, as opposed to that of the teacher as authority, respondents appear to be positively predisposed towards the idea. It should be noted at this point that enabling students to assume more responsibility for their own learning is an essential goal of DI and this can only be accomplished when the teacher relinquishes his/her role of classroom authority and acts as a guide and monitor (Smith and Throne, 2007).

However, a contradiction arises between this statement and the prevalence of whole class instruction. It is possible that respondents believe that they can take on different roles depending on the nature of the task at hand, i.e. that they can act as monitors at times and as controllers of activities at other times. Another explanation for this paradox may be that respondents agree with ceding a portion of their authority in principle, but, due to practical considerations, such as lack of time or pressure to cover the required materials, they may fall back on more familiar modes of instruction.

Lastly, it is also possible that some respondents may not have fully grasped what this change in teacher roles entails, accepting change on the surface, without fully understanding the principles and rationale behind it (Gardner, 2008). This, however, can have a detrimental effect on the implementation of differentiated techniques in Greek state primary schools. Incomplete or superficial understanding of new ideas in teaching has been shown to result in misunderstandings, confusion and teacher disillusionment, and, finally, to the abandonment of the recently introduced elements for the sake of more traditional practices (Karavas-Doukas, 1995, Li, 1998, Clark et al., 1999, Carless, 2003).

Differentiated Strategies

Turning to the issue of how frequently the various differentiated techniques are used, it would seem that some strategies are easier to adopt than others. Flexible grouping in pair/ group work activities appears to be more popular than differentiating reading texts. It is possible that this may be due to lack of training in the latter techniques or that practical matters e.g. lack of resources or increased preparation time may dissuade the teachers from putting these strategies into practice.

For instance, it should be noted that Greek state school teachers of English design their own paper-and-pencil tests. Producing a single effective, well-planned test is hard enough for most teachers. Designing multiple versions may indeed appear too demanding.

Furthermore, it has already been mentioned that the respondents' statements on individual work lead to the conclusion that it is still prevalent in Greek state school classrooms. Nevertheless, pair and group work appear to be popular as well. The majority of respondents claim to pair students with different levels of ability. Thus, struggling students are given the opportunity to learn from their peers and highly proficient students are motivated to cooperate with their classmates and consolidate their own knowledge in the process. On the other hand, they prefer to form groups according to the students' shared interests, so as each student can contribute what s/he can on a subject appealing to all group members.

On the whole, the findings indicate that differentiated techniques that require extra work and time for preparation on the part of the teachers are practiced less frequently than others that are viewed as more economical in terms of time and resources. This is consistent with Carless (1998), who stated that teachers' feeling stressed and over-burdened may lead them to develop negative attitudes towards the adoption of new practices, which may in turn result in the teachers' rejection of these practices for the sake of more familiar ones. Teachers should therefore be allowed the time and space to experiment with innovations, to reflect on potential implications for classroom practice and to appreciate their advantages (Rogers, 1995, Ellis, 1997).

Lastly, the findings on the effect of teachers' age and academic training seem to corroborate previous research suggesting that younger teachers and those with increased formal qualifications are more open to new ideas and practices, such as DI and differentiated strategies. This can have serious implications for the design of teacher training programs promoting the use of DI, as more time and effort may need to be devoted to helping more mature teachers and those with less academic training fully understand DI principles and incorporate differentiated techniques into their teaching repertoires.

6   Conclusion

The findings of the study indicate that Greek state school teachers of English acknowledge the increasing heterogeneity in classroom populations, as well as the need to design lessons with their students' unique characteristics in mind. They also foster positive attitudes towards DI in general and claim to adopt certain differentiated strategies e.g. flexible grouping. However, DI cannot be said to be fully implemented in Greek state schools and several aspects of it give rise to mixed feelings or negative reactions on the part of teachers, for example, the idea of differentiation of pace. Moreover, many participants do not appear to have fully grasped some of the changes that DI entails e.g. the change in teacher roles.

As a result, two suggestions could be made to materials developers and teacher trainers. Teachers may find training in this relatively new approach useful in order to clarify DI tenets in their minds and enrich their teaching repertoires with differentiated strategies. Teacher trainers need to take factors such as the teachers' age and academic training into account in order to design appropriate programs. Providing teachers with more differentiated materials and practical tips on how to implement DI more

effectively may also alleviate some of the burden of everyday preparation and persuade educators to incorporate more differentiated techniques into their teaching.

References

Avidov-Ungar, O. and Eshet-Alkakay, Y. (2011) Teachers in a World of Change: Teachers' Knowledge and Attitudes Towards the Implementation of Innovative Technologies in Schools. *Interdisciplinary Journal of E-Learning and Learning Objects*, 7, 291-303.

Bedee, S. (2010) *The Impact of Literature Circles on Reading Motivation and Comprehension for Students in a Second Grade Classroom*. Unpublished MA thesis. Graduate College of Bowling Green State University.

Benjamin, A. (2003) *Differentiated Instruction: A Guide for Elementary School Teachers*. Eye on Education, Columbus.

Burns, J. (2004) *An Analysis of the Implementation of Differentiated Instruction in a Middle School and High School and the Effects of Implementation on Curriculum Content and Student Achievement*. Unpublished PhD dissertation. Seton Hall University.

Carless, D.R. (1998) Review of Selected Factors Affecting the Implementation of Innovations. *System*. 26, 353-368.

Carless, D.R. (2003) Factors in the Implementation of Task-based Teaching in Primary Schools. System. 31, 485–500.

Clark, J., Lo, Y.C., Hui, M.F., Kam, M., Carless, D., Wong, P.M. (1999) *An Investigation into the Development and Implementation of the TOC Initiative With Special Reference to Professional Competencies*. Professional Development and Resources: Final Report. Hong Kong Institute of Education, Hong Kong.

Dörnyei, Z. (2003) *Questionnaires in Second Language Research*. Lawrence Erlbaum Associates, London.

Ellis, R. (1997) *SLA Research and Language Teaching*. Oxford University Press, Oxford.

Gardner, S. (2008) Changing Approaches to Teaching Grammar. *ELTED*. 11, 39-44.
Karavas-Doukas, E. (1995) Teacher Identified Factors Affecting the Implementation of an EFL Innovation in Greek Public Secondary Schools. *Language, Culture and Curriculum.* 8, 1, 53–68.

Kusano, K., Frederiksen, S., Jones, L., Kobayashi, M.. Mukoyama, Y., Yamagishi, T. et al. (2013) The Effects of ICT Environment on Teachers' Attitudes and Technology

Integration in Japan and the U.S. *Journal of Information Technology Education*, 12, 29-43.

Li, D. (1998) It's Always More Difficult Than you Plan and Imagine. Teachers' Perceived Difficulties in Introducing the Communicative Approach in South Korea. *TESOL Quarterly.* 32, 677–697.

Marulanda, M., Giraldo, P. and Lopez, L. (2006) *Differentiated Instruction for Bilingual Learners*. Presentation at Annual Conference of the Association for Supervision and Curriculum Development, San Francisco. 1-12.

McAdamis, S. (2001) Teachers Tailor their Instruction to Meet a Variety of Student Needs. *Journal of Staff Development*. 22, 2, 1-5.

McQuarrie, L., McRae, P. and Stack-Cutler, H. (2008) *Differentiated Instruction Provincial Research Review*. Alberta Initiative for School Improvement, Edmonton.

Ministry of Education, Lifelong Learning and Religious Affairs (2011) *Foreign Languages at School: Guide for the Foreign Languages Teacher*. Retrieved from http://rcel.enl.uoa.gr/xenesglosses/.

Rasmussen, F. (2006) *Differentiated Instruction as a Means for Improving Achievement as Measured by the American College Testing (ACT)*. Unpublished PhD Dissertation. Loyola University of Chicago, School of Education.

Rogers, E. M. (1995) *The Diffusion of Innovations* (3rd ed.). The Free Press, NY.

Smith, G.E. and Throne, S. (2007) *Differentiating Instruction with Technology for K-5 Classrooms,* ISTE (International Society for Technology in Education), New York.

Theisen, T. (2003) Differentiated Instruction in the Foreign Language Classroom: Meeting the Diverse Needs of All Learners. *Communique*. 6, 1-8.

Tieso, C. (2005) The Effects of Grouping Practices and Curricular Adjustments on Achievement. *Journal for the Education of the Gifted.* 29, 1, 60–89.

Tomlinson, C.A. (1999) Mapping a Route Toward Differentiated Instruction. *Educational Leadership*. 57, 1, 12-16.

Vastaki. M.S. (2010.) Η Διαφοροποίηση στη Διδασκαλία. *Επιστημονικό Βήμα. 12, January 2010*, 121-135.

Author Index