

**UNIVERSITY OF DEBRECEN**  
**KERPELY KÁLMÁN DOCTORAL SCHOOL**

Head of Doctoral School:  
Prof. Dr. János Nagy D.Sc.

Supervisor:  
Prof. Dr. János Nagy D.Sc.

**INTERNATIONAL COMPARATIVE ASSESSMENT OF  
SELECTED MID-TERM INSTITUTIONAL EVALUATION METHODS AND  
REGIONAL PROGRAMMES IN THE EU-12**

Prepared by:  
Balázs Mezősi  
Ph.D. candidate

Debrecen  
2015

INTERNATIONAL COMPARATIVE ASSESSMENT  
OF SELECTED MID-TERM INSTITUTIONAL EVALUATION METHODS  
AND REGIONAL PROGRAMMES IN THE EU-12

Dissertation to obtain PhD degree in Regional Sciences

Prepared by Balázs Mezősi M.Sc. in Economics

Kerpely Kálmán Doctoral School, Doctoral Programme in Regional Sciences

Supervisor: Prof. Dr. János Nagy DSc

<b>Final examination committee:</b>	name	degree
Chairman:	Prof. Dr. Sinóros-Szabó, Botond	DSc
Members:	Prof. Dr. Horváth, Gyula	DSc
	Prof. Dr. Stark, Antal	DSc

Date of final examination: 13 August 2015

<b>Dissertation opponents:</b>	name	degree	signature
Opponent:	_____	_____	_____
Opponent:	_____	_____	_____

<b>Dissertation committee:</b>	name	degree	signature
Chairman:	_____	_____	_____
Members:	_____	_____	_____
	_____	_____	_____
Secretary:	_____	_____	_____

Dissertation defense date: \_\_\_\_\_

# TABLE OF CONTENTS

1	INTRODUCTION	9
2	EVALUATION IN THEORY AND IN PRACTICE	16
2.1	<b>Background</b>	<b>16</b>
2.2	<b>Evaluation history in an interdisciplinary context</b>	<b>16</b>
2.2.1	A brief history of personal capability assessment	17
2.2.2	A brief history of evaluation in education	18
2.2.3	A brief history of aid programme evaluation	20
2.2.4	Summary on the historical, interdisciplinary overview of evaluation	22
2.3	<b>Evaluation theory</b>	<b>23</b>
2.3.1	Purpose of evaluation	23
2.3.2	Evaluation approaches	24
2.3.3	Common evaluation principles	29
2.4	<b>Programme evaluation in EU Cohesion Policy</b>	<b>30</b>
2.4.1	EU programme evaluation	30
2.4.2	Evolution of evaluation in the European Union	31
2.4.3	Programme evaluation subjects: interventions	34
2.4.4	Programming, implementation and evaluation cycle	36
2.4.5	Structure of Cohesion Policy interventions	37
2.4.6	Cohesion Policy programme evaluations in the 2007-13 period	39
2.4.7	Cohesion Policy evaluation methodology framework in the 2007-13 period	39
2.5	<b>Mid-term Cohesion Policy programme evaluations in 2007-13</b>	<b>41</b>
2.5.1	Evaluation context	42
2.5.2	Evaluation themes	42
2.5.3	Evaluation methods	45
2.6	<b>Formulation of research scope and hypotheses based on the summary of literature review</b>	<b>46</b>
2.6.1	Research scope	46
2.6.2	Research approach	47
2.6.3	Research hypotheses	48
3	INTERNATIONAL COMPARATIVE ASSESSMENT OF SELECTED MID-TERM INSTITUTIONAL EVALUATION METHODS AND REGIONAL PROGRAMMES IN THE EU-12	49
3.1	<b>Research set-up for the regional analysis and methodology analysis</b>	<b>49</b>
3.1.1	Research overview	49
3.1.2	Selection of research subjects for the methodology analysis	50
3.1.3	Countries, programmes and evaluations selected for methodology analysis	52
3.2	<b>Regional programme analysis: Cohesion Policy regional results in EU-12</b>	<b>55</b>
3.2.1	Cohesion Policy context in the EU-12	55
3.2.2	Cohesion Policy funds allocation country clustering in the EU-12	58
3.2.3	Regional impacts of Cohesion Policy in the EU-12 (2004-2011)	59

<b>3.3</b>	<b>Regional programme analysis: ROPs and ROP evaluations in the EU-12</b>	<b>63</b>
3.3.1	Regional development in the EU-12	63
3.3.2	ROP measure structure and content level comparison in the EU-12 countries	66
3.3.3	ROP mid-term evaluation analysis in the EU-12	68
<b>3.4</b>	<b>Methodology analysis: Simple Progress Overview Tool (SPOT)</b>	<b>73</b>
3.4.1	Method overview	73
3.4.2	International examples	80
3.4.3	Findings of the analysis	85
<b>3.5</b>	<b>Methodology analysis: Post contractual assessment (PCA)</b>	<b>87</b>
3.5.1	Method overview	87
3.5.2	International examples	91
3.5.3	Findings of the analysis	93
<b>3.6</b>	<b>Methodology analysis: Lead time analysis (LTA)</b>	<b>95</b>
3.6.1	Method overview	95
3.6.2	International examples	97
3.6.3	Findings of the analysis	104
<b>4</b>	<b>CONCLUSIONS</b>	<b>106</b>
<b>4.1</b>	<b>Thesis 1 – Cohesion Policy regional results</b>	<b>107</b>
<b>4.2</b>	<b>Thesis 2 – ROP diversity and similarity</b>	<b>108</b>
<b>4.3</b>	<b>Thesis 3 – Method objectivity, and link to intervention objectives</b>	<b>110</b>
<b>4.4</b>	<b>Thesis 4 – Method simplicity and robustness</b>	<b>112</b>
<b>4.5</b>	<b>Thesis 5 – Method flexibility</b>	<b>113</b>
<b>5</b>	<b>NEW SCIENTIFIC RESULTS</b>	<b>115</b>
<b>6</b>	<b>PRACTICAL APPLICATION OF RESULTS</b>	<b>117</b>
<b>7</b>	<b>SUMMARY</b>	<b>119</b>
<b>8</b>	<b>ÖSSZEFOGLALÁS</b>	<b>122</b>
<b>9</b>	<b>REFERENCES</b>	<b>125</b>
<b>10</b>	<b>PUBLICATIONS</b>	<b>134</b>
<b>11</b>	<b>ACKNOWLEDGEMENTS</b>	<b>136</b>
<b>12</b>	<b>DECLARATIONS</b>	<b>137</b>

## **GLOSSARY OF TERMS**

<b>ACIS</b>	Authority for Coordination of Structural Instruments, Romania
<b>AEA</b>	American Evaluation Association
<b>CEE</b>	Central and Eastern Europe
<b>CF</b>	Cohesion Fund
<b>CIP</b>	Competitiveness and Innovation Framework Programme
<b>COM</b>	European Commission
<b>CEE</b>	Central and Eastern Europe
<b>CPR</b>	Common Provisions Regulations (1083/2006 EC and the following 1303/2014 EC)
<b>CSF</b>	Community Support Framework
<b>DAC</b>	Development Assistance Committee (OECD)
<b>DG</b>	Directorate General
<b>EC</b>	European Commission
<b>EEA</b>	European Economic Area
<b>EES</b>	European Evaluation Society
<b>EMU</b>	European Monetary Union
<b>EQ</b>	Evaluation Question
<b>ERDF</b>	European Regional Development Fund
<b>ESF</b>	European Social Fund
<b>ESRI</b>	Economic and Social Research Institute
<b>EU</b>	European Union
<b>EU-2</b>	The two members of the European Union having acceded in 2007
<b>EU-10</b>	The ten members of the European Union having acceded in 2004
<b>EU-12</b>	The 12 members of the European Union acceded in 2004 and 2007
<b>EUR</b>	euro
<b>FP6/7</b>	Framework Program 6/7
<b>GDP</b>	Gross Domestic Product
<b>GBER</b>	General Block Exemption Regulation
<b>Hn</b>	Hypothesis (n=1 to 5, e.g. H1 stands for Hypothesis 1)
<b>IB</b>	Intermediary Body
<b>IPA</b>	Instrument for Pre-Accession Assistance
<b>IT</b>	Information Technology
<b>LTA</b>	Lead time assessment
<b>MA</b>	Managing Authority
<b>MECE</b>	mutually exclusive and collectively exhaustive
<b>NDP</b>	National Development Plan
<b>NSRF</b>	National Strategic Reference Framework
<b>OECD</b>	Organisation for Economic Co-operation and Development

<b>OECD DAC</b>	OECD Development Assistance Committee
<b>OP</b>	Operational Programme
<b>OP ACD</b>	Operational Programme for Administrative Capacity Development 2007-13 Romania
<b>OP TA</b>	Operational Programme for Technical Assistance 2007-13 Romania
<b>OPI</b>	Operational Programme I 2007-13 Malta
<b>OPII</b>	Operational Programme I 2007-13 Malta
<b>OPRD</b>	Operational Programme for Regional Development, 2007-13 Bulgaria
<b>PA</b>	Priority Axis
<b>PCA</b>	Post-Contractual Assessment
<b>POS CCE</b>	Sectoral Operational Programme for Improved Economic Competitiveness 2007-13 Romania
<b>PPCD</b>	Programming and Planning Department (Unit in Maltese Prime Minister's Office)
<b>QQTTP</b>	Indicator criteria set (Quantity, Quality, Time, Target group, Place)
<b>ROP</b>	Regional Operational Programme
<b>RTDI</b>	Research, technological development and innovation
<b>SF</b>	Structural Funds
<b>SI</b>	Structural Instruments (standing for ERDF, ESF and CF together)
<b>SMART</b>	Indicator criteria set (specific, measurable, achievable, relevant, time-bound)
<b>SME</b>	Small and Medium sized Enterprises
<b>SMIS</b>	The official monitoring and data management system in Romania
<b>SOP ENV</b>	Sectoral Operational Programme for Environment 2007-13 Romania
<b>SOP HRD</b>	Sectoral Operational Programme for Economic Development 2007-13 Romania
<b>SOP IEC</b>	Sectoral Operational Programme for Improved Economic Competitiveness 2007-13 Romania
<b>SOPT</b>	Sectoral Operational Programme for Transport 2007-13 Romania
<b>SPD</b>	Single Programming Document
<b>SPOT</b>	Simple Progress Overview Tool
<b>TA</b>	Technical Assistance
<b>ToR</b>	Terms of reference
<b>UK</b>	United Kingdom
<b>UMIS</b>	Unified Monitoring Information System (official monitoring system in Bulgaria)
<b>UN</b>	United Nations
<b>UN ERC</b>	United Nations Evaluation Resource Centre
<b>UNDP</b>	United Nations Development Program
<b>UNDP EG</b>	United Nations Development Program Evaluation Group
<b>WD</b>	Working Document
<b>WP</b>	Working Paper

## TABLE OF FIGURES

Figure 1 – Evaluation history by disciplines.....	17
Figure 2 – Aid evaluation history and overview.....	21
Figure 3 – Evaluation subjects comparison by discipline.....	22
Figure 4 – Evaluation approach comparison by discipline.....	22
Figure 5 – Main purposes of programme evaluation.....	23
Figure 6 – Evaluation and monitoring distinction.....	24
Figure 7 – Evaluation principles by determinant evaluation communities and their summary.....	28
Figure 8 – Evaluation main principles by determinant evaluation communities.....	29
Figure 9 – Development of evaluation activities over the programming periods from 1989 to 2020.....	34
Figure 10 – Programme and evaluation cycle of the Structural Instruments co-financed programmes.....	36
Figure 11 – NSRF-OP-PA-measure.....	38
Figure 12 – European Commission evaluation framework 2007-13.....	40
Figure 13 – Triangulation concept.....	41
Figure 14 – Evaluation activity framework.....	45
Figure 15 – Summary of research set up.....	50
Figure 16 – Summary of methods in the methodology analysis (focus, advantages, disadvantages).....	51
Figure 17 – Selection result overview table (by country, programme and evaluation) for H 3-5.....	53
Figure 18 – Selection result overview table by evaluation method (testing H3-5).....	54
Figure 19 – Overview of the EU-12 countries, basic information on GDP and EU funds.....	56
Figure 20 – 2007-13 EU-12 funds vs. GDP by country (2008 figures).....	57
Figure 21 – 2007-13 EU funds vs. population by country (2008 figures).....	57
Figure 22 – 2007-13 EU funds breakdown by country in the CEE.....	58
Figure 23 – Clusters of EU countries – GDP per capita vs. EU funds in 2007-13 per capita.....	58
Figure 24 – EU-12 GDP/capita figures 2004-2014.....	60
Figure 25 – EU-12 GDP/capita figures 2004-2011 at NUTS2 level (EU-28 = 100).....	61
Figure 26 – EU-12 NUTS2 level development gap figures 2004-2011 (EU-28 = 100).....	62
Figure 27 – EU-12 country regional operational programme and intervention details.....	64
Figure 28 – Intervention structure and content of EU-12 regional development measures.....	67
Figure 29 – Focuses of evaluation stages in the 2007-13 programming period.....	69
Figure 30 – Summary of analysis findings (ROPs vs. EC Working Document No.1).....	72
Figure 31 – A standard project selection (application) process using a three-stage selection system.....	74
Figure 32 – Breakdown of the three-stage selection system to processing and approval elements.....	75
Figure 33 – SPOT equation tool (assuming a three-stage project selection system).....	76
Figure 34 – SPOT equation tool (on a three-stage project selection system) explanations for low ratios.....	79
Figure 35 – Progress factors at 30/06/2009 and 30/06/2010 by OP of the Romanian NSRF (%).....	80
Figure 36 – SPOT equation of SOP IEC at 30.09.2010.....	81
Figure 37 – SPOT equation of OPI Malta ERDF measures at 30.09.2010.....	82
Figure 38 – SPOT equation of OPI Malta CF measures at 30.09.2010.....	82
Figure 39 – SPOT factors of progress of OPRD Bulgaria at 31.12.2010.....	84
Figure 40 – SPOT evaluation findings by hypothesis (H3-5).....	85
Figure 41 – Planned vs. Actual Contracting / Disbursement Schedule in OPI, Malta evaluation PCA.....	88
Figure 42 – Planned vs. Actual Expenditure in OPI, Malta evaluation, PCA.....	89
Figure 43 – Planned vs. Actual Implementation Schedule in OPI, Malta evaluation PCA.....	90
Figure 44 – Stage specific lags, summary of results in OPI, Malta evaluation PCA.....	91
Figure 45 – Summary of stage specific lag estimates in OPI, Malta evaluation PCA.....	92

Figure 46 – Summary of project specific lag estimates in OPI, Malta evaluation PCA .....	92
Figure 47 – PCA evaluation findings by hypothesis (H3-5).....	93
Figure 48 – A standard project selection (application) process using a three-stage selection system .....	96
Figure 49 – Lead time analysis standard data collection sheet .....	96
Figure 50 – Summary table of lead times in OPRD, Bulgaria evaluation .....	98
Figure 51 – Lead time analysis from registration to technical and financial check OPRD, Bulgaria.....	98
Figure 52 – Lead time analysis rolling and set deadline procedures OPRD, Bulgaria evaluation .....	99
Figure 53 – Lead time analysis from technical and financial check to approval OPRD, Bulgaria .....	100
Figure 54 – Lead time analysis from registration to contracting results OPRD, Bulgaria evaluation .....	100
Figure 55 – Lead time analysis of the application process by calls OPII, Malta evaluation .....	103
Figure 56 – Lead time analysis of the application process by schemes (MA and IB) OPII, Malta .....	103
Figure 57 – LTA evaluation findings by hypothesis (H3-5).....	104
Figure 58 – Summary table of conclusions by hypothesis (H3-5).....	106

# 1 INTRODUCTION

## **Relevance and rationale of the topic**

Evaluation of funding programmes is a natural expectation of all stakeholders (donors, acceptors, policy makers, broad public), driven by the absence of institutionalised feedback mechanisms, such as the market itself in market economy context.

Evaluations require methods and tools to provide answers to stakeholder questions. Among the many other existing classifications, these methods can be objective or subjective in terms of the outputs they provide.

Evaluation methods that are flexible enough to yield objective, comparable and repeatable, precise outputs are the ones that provide unique opportunity to be used in cross-country cases or even at the entire Community level. Besides, methods are usually required to be simple to carry out and communicate and robust. These two expectations are seemingly contradictory, however, there are methods capable of balancing between the two. Actually, it is a logical request of evaluators, evaluation commissioners and above all, the final users of evaluations, the policy makers to have such methods and tools.

Public interventions in general – and EU funded programmes in particular – have three objectives: effectiveness (and relevance), efficiency and accountability (MEANS, 2000). As a result, evaluations most often concentrate on these topics, as evaluation themes.

In summary, policy makers and programmers need evaluations (and methods that back them up) as inputs to their activity, that on the one hand resolve the contradiction between broad use and simplicity, and also address objectivity and each of the three above mentioned objectives.

In this research, I picked three mid-term evaluation methods that I have developed, tested and applied in several EU countries, and I aimed at presenting and analysing their application in international comparisons. These methods have been used in many EU mid-term programme evaluations in the 2007-13 period, out of which I chose five evaluations from three countries for my research to observe more closely – Bulgaria, Malta and Romania.

In order to place the previously mentioned evaluation methods into context, an analysis of EU-12 Cohesion Policy intervention results, EU-12 regional (operational) programmes and their evaluations form part of my research.

As a result, my research is an international comparative analysis of EU mid-term evaluation methodologies, backed up by a Cohesion Policy programme results and regional programme analysis, addressing five hypotheses.

My first hypothesis assumes that EU Cohesion Policy could contribute to the development of regions in the EU-12, however, it could not counteract the growing interregional disparities in these countries. My second hypothesis addresses the diversity and similarity of Regional Operational Programmes in the EU-12 in the 2007-13 period. The last three of my hypotheses are based on assumptions that the selected evaluation methods fulfil the requirements of objectivity, simplicity and robustness, flexibility and addressing evaluation themes driven by intervention objectives (relevance and effectiveness, efficiency and accountability).

I am aware that my topic selection is on the verge of regional development sciences and management sciences. However, the fact that the selected methodologies have been widely used throughout Europe in evaluations in regional development programme context, creating the foundations for future programming, makes them relevant driving factors of regional development. As such, they make appropriate subjects for scientific research within the scope of regional sciences.

### **Personal motivation**

I have been evaluating EU programmes and policy documents since 2003. I had the privilege to be among the few experts who had the opportunity to take part in each and every evaluation stage (i.e. ex ante, mid-term and ex post) of the Structural Funds and Cohesion Funds co-financed Hungarian programmes in both past EU budgetary periods (i.e. 2004-06, 2007-13) and I have also been involved in the ex ante phase of the 2014-20 period. I carried out evaluations in six Central and Eastern European countries (Bulgaria, the Czech Republic, Estonia, Hungary, Romania and Slovakia) and Malta, commissioned by the national managing authorities or the respective Directorate General of the European Commission.

I have been project manager, and lead methodological expert and methodology developer in programme evaluation activities for the last twelve years. I had the opportunity to develop, challenge and test methodologies, tools and techniques to support evaluation tasks. Also, I have been member of the European Evaluation Society since 2010.

### **Scientific relevance and purpose**

There have been methodology handbooks issued by both the EU Commission (DG Regio), member state central and local governments, professional organisations such as associations and other stakeholders on the practical means of programme evaluation. I will cover and refer to the most relevant of these works in this dissertation as theoretical and scientific bases supporting real, practical evaluation activities. What is scarce, however, is new methods, practical feedback and guidance on using evaluation methods in diverse context. The feedback I am reflecting in this study has evolved through 20+ programme evaluations in 7 countries over 10+ years, capable of acting as impartial and focused synthesis to enrich and complement the scientific background already available on evaluation methodology. Apart from the attempt to act as a small building block to science, I also aim to provide evaluators with practical advice on what to use, and how to use when planning evaluation tasks.

The objective of my research is to thoroughly assess the three selected methodologies used for mid-term evaluations applied in EU Member states' National Strategic Frameworks and Operational Programmes in the 2007-13 EU budgetary period. By providing a review of the selected evaluation methods and mid-term evaluations, I aim to assist the evaluators and programmers of the 2014-20 period by supporting their choice of methods and tools; and also to share lessons learnt on the potential pitfalls of their application. From a scientific perspective, this dissertation serves the following purposes:

- 1.** Directly: Analysing internationally validated and accepted EU programmes, mid-term evaluations and evaluation methods of the 2007-13 programming period in the EU-12
  - i) Presenting the high level regional results of Cohesion Policy in EU-12;
  - ii) Providing a detailed description and analysis of the methodologies;
  - iii) Demonstrating their use through international examples;
  - iv) Synthesising the advantages and limitations of the application of methods.

2. Indirectly: Supporting evaluation activities foreseen for the 2014-20 programming period by formulating conclusions on existing methodologies and focuses:
  - i) Supporting EU Cohesion Policy related programme development and evaluation;
  - ii) Supporting non-EU programme planning;
  - iii) Facilitating evaluation capacity building on both programmers' and evaluators' side.

### **Scope**

My research is limited to the EU-12 countries, consisting of the EU-10 countries entering the Community in 2004 and the EU-2 (i.e. Romania and Bulgaria) accessing in 2007<sup>1</sup>.

The scope of my research is three well-acknowledged mid-term programme evaluation methodologies applied in five mid-term evaluations of National Strategic Reference Frameworks (NSRFs) and Operational Programmes (OPs) in 2007-13 in three EU-12 countries: Bulgaria, Malta, and Romania, with an underlying outlook to EU-12 EU co-funded regional development programmes and regional development measures and their results.

The driving factor for the choice of the *mid-term evaluation phase* methods in my research scope lies in three points. First, mid-term evaluations carry strong view on development policy, and within that a sharp focus on the institutional aspect, i.e. relevance and rationale, efficiency and effectiveness. Second, mid-term evaluations cover the most by assessing the performance of the actual period retrospectively while they also formulate recommendations for the next programming period. Third, 2007-13 mid-term evaluations are the latest relevant (there are 2014-20 ex ante evaluations available, but their goal, scope and therefore their choice of method is different) approved evaluations available. 2007-13 ex post evaluations are just to be commenced in member states, however, their character is much different, concentrating on the financial and administrative aspects of the development programmes.

<sup>1</sup> The „EU-12” originally referred to the first 12 members states of the EU, however, nowadays more and more official reports refer to the EU-10 (accession in 2004) + EU-2 (accession in 2007) countries as „EU-12”, also including the Sixth Cohesion Report of the EU.

Besides, my research covers many aspects of EU programmes and the underlying EU Cohesion Policy and its intervention structure. It also covers evaluation, especially mid-term evaluation of such programmes.

I do not aim to give a full review of the selected programmes, and evaluation methods, tools or techniques. I do not aim to cover all the evaluation methods, tools and techniques within these or other evaluations. I picked three evaluation methods that have proven in many evaluations to be capable of potentially providing answers for a large proportion of evaluation questions of an evaluation theme in the course of a mid-term evaluation. In general, the application of these methods have proved to account for the majority of the findings, conclusions and recommendations in the institutional aspect of mid-term evaluations I previously conducted.

Besides, this study carries important conclusions on how programmes and evaluations, in which these methods have been used, had been set up in general, i.e. what regional development approach they had followed and what they had focused on.

The research covers only programmes that have been co-financed by EU Structural Funds (SF) and Cohesion Funds (CF) (together: Structural Instruments (SI)) available in the 2007-13 period. Norway Grants, EEA Grants, Sixth and Seventh Framework Programs (FP6, FP7), Competitiveness and Innovation Framework Programme (CIP) do not form part of the research.

The research does not explicitly cover 2004-06 and 2014-20 programmes, though it does include some references to these EU budgetary periods and the respective programmes. In terms of evaluations, my research marginally covers non-mid-term (e.g. ex ante, ex post) evaluations required merely to act as reference points to position mid-term evaluations in terms of timing.

### **Tasks and methodologies**

I planned and executed the research as an iterative process which finally led to the findings, lessons learnt, conclusions and new scientific and practical results.

The primary means of testing the above hypotheses was desktop research and literature review. This activity covered the relevant regulations, policy documents, methodological papers, programmes and their evaluations and other related documentation. I also

conducted interviews with relevant programme managers and programme evaluators in all countries I researched to cross-check and challenge the hypotheses in the course of methodology development, the results of which have been used as input to this study.

In the current research, I used multiple data collection and analysis techniques. The primary objective was the systematic information collection, covering programming documents (NSRF, OP, Action Plans, measures, etc.), relevant (mid-term programme) evaluations, a sample of / extract from application files, strategic and monitoring reports, status reports on the implementation, annual reports, published, aggregated progress tables for validation of project level data set and other background strategic documents.

Data analysis included the examination of the availability, the relevance and the suitability of the indicators in order to map intervention logic, to take hold of and operationalise the results as well as to follow up development (baseline, milestone, actual indicator values). It also covered an analysis of the data collected from the implementation system and other databases, from statistical and research data and collected data, e.g.: project level data sets (UMIS, SMIS), chronology of calls request, the financial allocation by year by intervention or operation and the publicly available statistical data used from Eurostat and national statistical offices.

I also conducted personal interviews with the stakeholders, evaluators and policy officers directly or indirectly affected by the elaboration, implementation or results of the programmes; exploring their opinion and their attitude to the results, documenting the collected information in detail.

### **Structure of the dissertation**

The dissertation consists of three main sections, the first being the present introductory section.

In the second section, I aim to provide an overview of the theoretical, scientific and practical background and evolution of evaluation. My research has been built on the foundations provided by the processing of literature, consisting of evaluation history (inter alia based on Patton, 1997; Legge, 1983; Horváth, 2015; Pálné, 2009; Epstein and Prak, 2010; Gilbert, 2010; Sasaki, 2006; Hirschmann, 1968; Rossi, Freeman, and Wright, 1979; Cassen, 1986; UN and OECD documentation) and evaluation theory (based inter alia on Weiss, 1997; Huey-tsyh, Chen, 2005; Sanders, F. – Worthen, 2004; OECD, 1991;

OECD, 2008; UNDP, 2009; AEA, 2004, Tavistock, 2003), evaluation approach and classifications (based on Potter, C., 2006; Pollitt, C.–Summa H., 1996; Pawson, R. - Tilley, N., 1997; Scriven, M., 1991, Weiss, 1999; Louw, 1999; Molle, 2007; Jobbágy, 2014; Nagy, 2005 and Lóránd, 2009) intervention logic and classification (based on Fitz Gerland, 1999; ESRI, 1993; EC, 2006).

This section is followed by a general overview of practical EU evaluation context, encompassing EU evaluation related regulations (EC, 1998, EC, 1999a-g, EC 2006g-k, EC, 2014a-b) and methodology guidance (based on Scriven, 1991; Potter C, 2006; EC, 2000a-j; EC, 2006a-f; MEANS, 2000; OECD, 2002; Tavistock, 2003), complemented by EU evaluation framework (EC, 2007; EC2006a-f; UNDP, 2009, MEANS, 2000) and evaluation commissioning experience in the selected countries of Bulgaria, Romania and Malta (based on ACIS 2010; UMIS, 2011, PPCD, 2011; SOP IEC MA, 2010).

In the third section, I present the methodology analysis: methodologies used in mid-term programme evaluations in the selected countries in 2007-13 and provide an international comparison and assessment on their application. (ACIS, 2010; PPCD, 2011, MRDPW, 2011, UMIS, 2011). Also in this section, I present a regional programme analysis on the Regional Operational Programmes of the EU-12 countries selected for the research, with special focus on, Bulgaria, Malta and Romania (EC, 2008; Eurostat, 2008)<sup>2</sup>.

The dissertation is closed by research conclusions, new scientific results and practical application of the results, a brief summary of my dissertation in English and Hungarian and the references and publications sections.

<sup>2</sup> The reason for using 2008 data is that reference data should comply with the 2007-13 period regional policy intervention values set in 2008.

## **2 EVALUATION IN THEORY AND IN PRACTICE**

As my research has the assessment of evaluation methodologies in its focus, the first part of the current section contains an interdisciplinary overview on the history of the development of the evaluation discipline. I approach the dissertation focus topic step by step by assessing evaluation theory in general, then moving on to programme evaluation in the EU and eventually arriving at mid-term programme evaluations (covering most of my evaluation topics and this way the later analysed methodologies as well) in the 2007-13 programming period. This logical path and the pieces of information linked to it will act as pillars to the theoretical foundation of defining the hypotheses (see Section 2.6.3 for more detail) to my research on methodologies.

### **2.1 Background**

Evaluation is a common, everyday activity. “Evaluation is an elastic word that stretches to cover judgements of many kinds” (Weiss, 1997). Its meaning covers inter alia “examination”, “test”, “assessment”, “appraisal”, “diagnosis” and “review”. Based on this list of notions, it is obvious that evaluation is an activity, therefore it is a mean to and end and not an end itself (Patton, 1997).

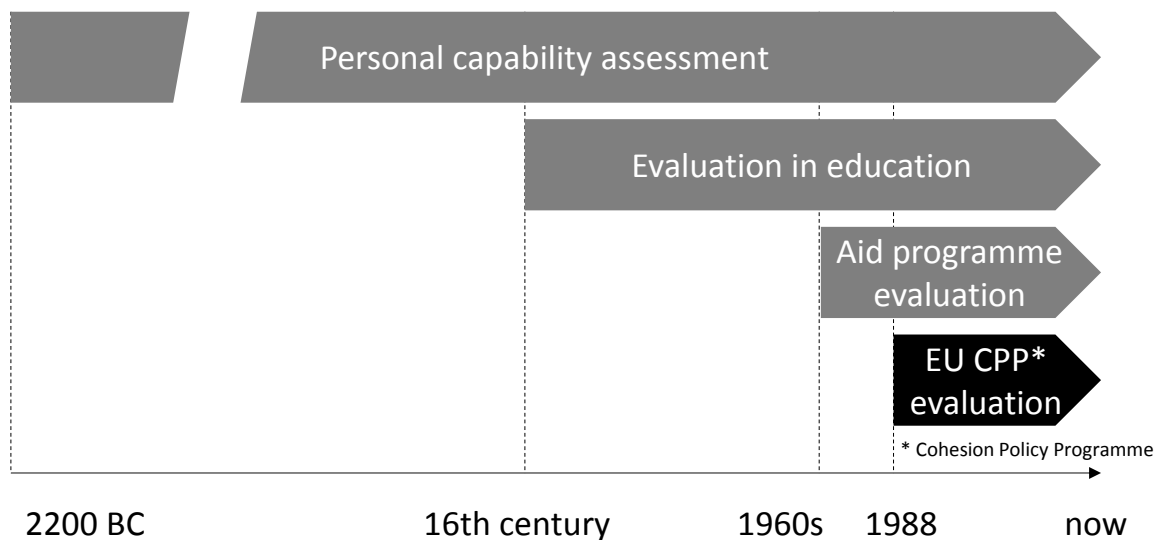
Evaluation is an element of science, a discipline, primarily that of social sciences. The roots of evaluation come from human resource management, education and politics, and have gone through a long evolution up till today, when it is still education, personnel performance, policy and funding programs where it is most relevant. Evaluation has become a focus area with policy making and funding programmes in the late 20<sup>th</sup> century, and still maintain to be acknowledged as a means to inform policy makers, fund donors and the broad public on what has been achieved, why and how.

### **2.2 Evaluation history in an interdisciplinary context**

Evaluation nowadays is present in many academic disciplines and many fields of the socio-economic environment. It has more than four thousand years of history, with two major foundations to the present evaluation science, one being personal capacity assessment, dating back to ancient and medieval times, the other being evaluation in education.

In the next section I follow a historical perspective, pursuing an interdisciplinary way how evaluation has evolved. In a chronological order, personal capacity (also including competence, ability, capacity) assessment related occurrences is followed by educational assessment (examinations, oral and written tests), closing on to aid programme (funding programme) evaluation and as part of that, the focus of my dissertation: evaluation of Cohesion Policy programmes (see Figure below) to be later specified to mid-term Cohesion Policy programme evaluations and their respective methods.

**Figure 1 – Evaluation history by disciplines**



Source: Own construction

### 2.2.1 A brief history of personal capability assessment

Institutionalised evaluation of personal capabilities has a 4500 year history, the first occurrences dating back to the Ancient Chinese and Ancient Greek times. In terms of activities, the roots of evaluation originally came from the assessment for approval to public positions for high level state and province officials. The reason behind this early type of evaluation was “quality control” on the human resource in the operation of public administration offices. These systems usually complemented (and ran parallel to) the inheritance system in which noble families could take such positions based on right of birth. Two prominent examples on the origins of personal capability assessment type of evaluation are the “imperial examinations” and the “dokimasia”.

The first appearance of evaluation date back to ancient China when Empire Shun (2256-2205 BC) required officials to go through an annual frequency competence test, and based on the results they were either promoted or dismissed from their duties (Legge, 1983).

After the resignation of Shun the system has slowly faded away. Much later, the imperial examination system for civil service has been applied by and since the reign the Sui Dynasty (581-618 AD). The imperial examination system for evaluating and selecting officials was based solely on merits (Zheng, 1994).

In Greece, *dokimasia* was used as a form of candidate evaluation at the city of Athens to acknowledge the capacity of the citizens for exercising public rights and duties (Adeleye, 1984). This system has been later adopted by the Roman Empire as well. The notion of „docimology” the discipline covering student evaluation originates from this ancient Greek word, made up from „dokime” (i.e. “attempting” in Greek) and „logos” (i.e. “science” in Greek).

Later on, in the Medieval Ages public offices were generally taken by noblemen as a birth right. However, craftsmen and artisans started gathering into guilds in the early 13<sup>th</sup> century to defend the quality and prices of their work. All who wished to join the guild of a certain craft, previously had to train alongside a master, assisting and learning at the same time. After learning the ins-and-outs of the craft, the apprentice broadened his views with different masters, preferably abroad. Admittance to a guild was linked to the creation of a masterpiece, a high quality, self-forged product that was capable of showing that the apprentice has mastered the craft (Epstein, Prak, 2010). In case of a positive result of this personal capability assessment, the candidate was accepted in the ranks of the guild.

In the 20<sup>th</sup> century, with the swift development of the disciplines of psychology, new approaches, methods and tools arisen to test people. They have been used in various cases: recruiting employees, performance evaluation in workplaces, defining target groups for marketing activities or even profiling criminals. Interviews, focus groups, tests and tools are widely used nowadays in learning more about people and their expectable performance and behaviour.

### **2.2.2 A brief history of evaluation in education**

Despite its distant roots, evaluation in its current frameworks is a relatively new phenomenon in education. The review on the results of the knowledge transferred to students did not inherently form part of the standard education processes. The oral examination of students became part of the curricula only in the late 16<sup>th</sup> century.

As for written exams, the first proven occasion dates back to 1702, the Trinity College of Cambridge (Ball, 1889).

In the 1840s, Horace Mann struggled for the wide acceptance of written examinations through setting objective and standard criteria for all students equally. The need for such systems have been underpinned by many education experts, however, scientific literature refers to Joseph M. Rice as the founder of examinations and testing, who designed this method in 1897. The word “test” occurs first in the publication of James M. K. Catter in a 1890 study. He was scientific assistant to Wundt, the famous psychologist and linguist, therefore the first tests were means of psychology. However, the discipline of docimology (originating from dokimasia) was launched by Henri Pieron, French psychologist, who elaborated the discipline of evaluation and performance grading through scientific methods (Gilbert, 2010).

According to the Columbia University, educational measurement dates back to 1904, when education psychologist Edward L. Thorndike published “An Introduction to the Theory of Mental and Social Measurements”, since quoted as the first textbook to define the knowledge base of classical test theory. During the 1950s, the “Cognitive Ability Tests” were widely used to test scholastic ability. In 1971, Thorndike and Hagen, the creators of the test, also co-edited the second edition of Educational Measurement, which has since become the best-known reference handbook in the field (Chatterji, 2013). According to Hallie Preskill and Darlene Russ-Eft (2005), evaluation showed accelerated evolution in the late last century. However, in the United States in 1950’s and 1960’s, evaluation still meant educational assessment, carried out by social science researchers in certain universities and organizations.

Educational evaluations can be categorised to programme, personnel and student evaluations. Programme evaluations concentrate on the knowledge transfer and its means and techniques, personnel evaluation is concerned with the capabilities and competencies of the knowledge sharer, while student evaluation is the classic assessment on the knowledge acquired. The purpose of the main focus area, i.e. student evaluation in education today is manifold: to discover the extent of competence and certify the level of knowledge, to appraise the status of and change of behaviour, to diagnose strengths and weaknesses and to assess progress made, to locate areas where remedial action is required, and to motivate for better attainment and growth.

### 2.2.3 A brief history of aid programme evaluation

The history of evaluation from 1940s to 1990s is based on the study conducted by Ryoh Sasaki (Sasaki, 2006) providing a concise outlook on the evolution of evaluation. With the establishment of World Bank in 1944 and agencies of the U.N. in 1945, aid programmes have been launched. The first evaluation in this time was the Measures for the Economic Development of Under-Developed Countries by Lewis (UN, 1951). In this period, aid solely focused on post WWII reconstruction and the re-establishment of infrastructures and networks (transport, electricity, etc.).

In the beginning of the cold war era, the United States and the Soviet Union competed on the platform of aids to third world countries as well. For these project aids and food aids, the Evaluating Development Projects was prepared for UNESCO. This was the first occasion of using the expression “evaluation” officially.

In the 1960s, many bilateral agencies came into existence in Europe and overseas. In the 1960s, cost-benefit and cost-efficiency analyses were emerging. Development Project Observed (Hirschman, 1968) was considered as the pioneer of evaluating aid projects. It is important to note that according to the author, the impacts of the aid projects were generally disappointing. Another notable event of mid-1960s is that The War on Poverty and the Great Society programs required large investments in the US, and Senator Robert Kennedy delayed their approval until the submission of an evaluation plan by local education agencies, and summary reports by state agencies. As a result, evaluation requirements joined to federal grants.

In the 1970s, aid focus shifted to agricultural and social development, in line with an ideological shift from economic development to reduction of poverty. Agencies conducted evaluation on aid projects, focusing on design and implementation. However, methods tried and tested in the economic development and infrastructure aid projects did not work in the social sector. As a result, Rossi, Freeman, and Wright were invited to present their ideas on evaluation research, which were later used as a basis for the milestone book *Evaluation: A systematic Approach* (Rossi, Freeman, and Wright, 1979).

In the 1980s, market based reform dragged along aid activities. In this period, in line with a distrust in governmental organisation and shifting toward market, the effectiveness and impact of governmental aid activities was questioned. Evaluation methods used in this

period were summarised in a study, with the title *Does Aid Work?* (Cassen, 1986) The new methodological developments in aid evaluation have been published as *Methods and Procedures in Aid Evaluation: A Compendium of Donor Practice and Experience* (OECD, 1986). The EU developed its own system, the Project Cycle Management (PCM) to use for evaluation. PCM was widely acknowledged and preferred by European donors, but also in Japan since then.

In the 1990's government program accountability, lean, efficient, global and more competitive organisations were required by the broad public. As a result, evaluation became commonly used in government mandates, but also to improve program effectiveness, enhance organizational learning, and inform allocation decisions in a wide variety of both public and private organizations. In the first decade of the new millennium, interest moved towards participatory, collaborative and learning-oriented evaluations and national evaluation associations have been established all around the world. The OECD-DAC Working Party on Aid Evaluation approved the DAC five evaluation criteria (i.e. relevance, effectiveness, efficiency, impact, and sustainability), which later became the basis of European Union evaluation perspectives. The next Figure (from Sasaki, 2006 based on Hjertholm, P & White, H., 2000) is an excellent and transparent summary and comparison of aid programmes in between 1940 and 1990:

**Figure 2 – Aid evaluation history and overview**

*Review of main developments in the history of foreign aid and its evaluation activities*

	Dominant or rising institutions	Donor ideology	Donor focus	Types of aid	Trends in aid evaluation	Aid evaluation reports / Influential books
1940s	Marshall Plan and UN system (including World Bank).	Planning.	Reconstruction.	- Marshall Plan (largely program aid)		
1950s	United States, with Soviet Union gaining importance from 1956.	Anti-communist, but with role for the state.	Community Development Movement.	- Food aid - Projects aid		Lewis (1951) Hayes (1959)
1960s	Establishment of bilateral programs	As for the 1950s, with support for state in productive sectors.	Productive sectors (e.g. support to the green revolution) and infrastructure.	- Bilaterals gave TA & budget support - Multilaterals supported projects.	(The boom in establishment of bilateral aid agencies.) - <i>Sophistication of economic analysis for project preparation</i>	Campbell & Stanley (1966) Hirschman (1967)
1970s	Expansion of multilaterals especially World Bank, IMF and Arab-funded agencies).	Continued support for state activities in productive activities and meeting basic needs.	Poverty, taken as agriculture and basic needs (social sectors).	- Fall in food aid - Start of import support.	The boom in establishment of evaluation units. - <i>Focus on the process of aid</i> - <i>Large-scale experiments &amp; disappointment.</i> - <i>Log-frame by USAID</i>	Tendler (1975) Rossi, Freeman & Wright (1978)
1980s	Rise of NGOs from mid-1980s.	Market-based adjustment (rolling back the state).	Macroeconomic reform.	- Financial program aid - Debt relief.	Focus on aid effectiveness by various approaches - <i>Rapid, low-cost methods</i> - <i>Participatory approach</i> - <i>Training &amp; empowerment</i>	Scriven (1980) Casley & Lury (1982) Cassen (1986) OECD (1986) Riddell (1987) Chambers (1988)
1990s- (Late '90s – 2000s)	Eastern Europe & FSU become recipients; emergence of corresponding institutions. Aid coordination forum at local level as well as headquarter level	Market-based adjustment continued Move back to role of the state. (Balance the market & the state)	Environment and gender (but passed quickly). Poverty and then governance	Financial program aid & debt relief continue Move toward sector support.	Diffusion of DAC's evaluation criteria vs. Adoption of "Managing results" approach. - <i>Sector level, country level, and thematic evaluation.</i> New agenda: Sector program evaluation New agenda: Evidence-based evaluation in aid	OECD-DAC (1991) " <i>DAC evaluation criteria</i> " Osborne & Gaebler (1993) World Bank (1992) " <i>Wapenhans report</i> " World Bank (1998) Wolfenson. (1998) Hatry (1999)

*Note.* Entries are main features or main changes, there are of course exceptions.

Source: Sasaki, 2006

## 2.2.4 Summary on the historical, interdisciplinary overview of evaluation

The three origins (i.e. personal capability assessments, educational and aid evaluation), of the discipline we call evaluation today bear many similarities and differences at the same time. From a historical point of view, personal related evaluations (also including personal related evaluations in education) have been complemented by programme related evaluations just a few decades ago. Both areas show different evolution running in parallel. Programme related evaluations only have cca. 80 years of history, compared to the 4500 years of personnel related evaluations. All three types of evaluation concentrate on the three main aspects, i.e. *the donor, the acceptor and the vehicle* of the activity (see Figure 3 below). However, in personal capability assessments, and educational evaluation, the focus is clearly on candidate evaluation – with examiner/ teacher performance evaluations gaining more space in democratic, transparent contexts. Aid evaluation, on the contrary, is rarely concentrating on the acceptor, i.e. aid beneficiary, it usually addresses the delivery mechanisms and the capacities of the stakeholders included, i.e. institutional system for aid delivery and programme performance.

**Figure 3 – Evaluation subjects comparison by discipline**

Aspects	Personal capability assessment	Educational evaluation	Aid evaluation
Donor	Examiner performance	Lecturer performance	Institutional system
Acceptor	Candidate performance	Student performance	Aid beneficiary performance
Vehicle	Acceptance process	Educational programme	Programme performance

Source: Own construction

Using the summative-formative evaluation distinction (based on MEANS, 2000, see Section 2.4.1) personal capability assessment and education evaluation are most *often summative* (taking account of a current snapshot of performance) and use qualitative methods, as opposed to aid evaluation, which is most *often formative* (with an approach to further development and continuous improvement) using extensively qualitative methods along with quantitative ones.

**Figure 4 – Evaluation approach comparison by discipline**

Aspects	Personal capability assessment	Educational evaluation	Aid evaluation
Evaluation type	Summative	Summative / Formative	Formative / Summative
Methods used	Qualitative	Quantitative / Qualitative	Quantitative / Qualitative

Source: Own construction

## 2.3 Evaluation theory

In the public sphere, evaluation is a tool (similarly to monitoring and audit) used in management of public policies, programmes and projects. According to the EU central administration definition “A judgement of interventions according to the results, impacts and needs they aim to satisfy” (EU Commission) and “A critical and detached look at the objectives and how they are being met.” (UK Treasury).

### 2.3.1 Purpose of evaluation

There is no such direct feedback mechanism in the public sector as the market in the private sector, which indicates unambiguously on the success of an intervention (Sanders, Worthen, 2004), therefore, it is evaluation that plays this role.

I can identify myself with the EC definition which states that the purpose of evaluation is an “Overarching learning purpose” and in the EU programme evaluation context it is “to learn through systematic enquiry how to better design, implement and deliver public programmes and policies” (MEANS, 2000). Main purposes of evaluation in the intervention context and the corresponding timing along the policy cycle are the following (the timing aspect is covered in more detail in Section 2.4.4):

**Figure 5 – Main purposes of programme evaluation**

Focus	Purpose of the evaluation	Timing
Planning	Supporting allocation of resources	Intervention start (Ex ante)
Implementation	Improving programme performance	During intervention (Mid-term / ongoing)
Accountability	Controlling compliance	Intervention close (Ex post)

*Source: Tavistock, 2003, own edition*

Besides, evaluation is sometimes used for other purposes and hidden agendas, such as justification of decisions already made, or a means to postpone decisions. Evaluations are often looked upon and used as means of political public relations and ways of ensuring mandatory compliance in a formal way. “A rational exercise often undertaken for non-rational reasons” (Weiss, 1997). In the current study I do not follow up on these latter instances of purpose, but follow the profound assumption of “an overarching learning purpose” (MEANS, 2000).

It is important to highlight at this point the differences between evaluation and monitoring. Monitoring and evaluation of programmes both cover – inter alia - means to keep track of progress along the pre-set programme objectives. Though they are generally linked processes, however, there are some fundamental differences. While monitoring is an ongoing activity, evaluation is discrete. Evaluators use information and indicators collected by monitoring. Monitoring fulfils the role of continuously informing decision makers on programme implementation, including an early warning system and highlighting areas that potentially require evaluation (ÁSZ, 2007).

**Figure 6 – Evaluation and monitoring distinction**

Aspect	Monitoring	Evaluation
<b>Goal</b>	Adjustment	Learning
<b>Level</b>	Operational	Strategic
<b>Focus</b>	Performance	Reasons behind findings, impacts
<b>Agent</b>	Internal	External
<b>Duration</b>	Continuous	Discrete

*Source: own construction based on own work experience*

### 2.3.2 Evaluation approaches

Apart from knowing *why* evaluate programmes (purpose), it is also important to identify *how* to evaluate programmes (approach) according to existing scientific resources and practicing organisations.

There are many schools and philosophies on evaluation approach. In order to provide a balanced look and a ground for critical overview, I cover both scientific literature sources and professional sources to show and challenge how evaluation should be carried out, what underlying principles and approaches exist.

#### Scientific literature review

There are many scientific approaches available on evaluation. In this sub-section I provide an overview, ranging from the academic to official EC guidance on evaluation approaches as they have a determinant role on the choice of applied evaluation methods and tools.

Starting with to the Guide (Tavistock, 2003), there are five different methodological approaches to evaluation of interventions. As the Guide has been compiled as a results of

great scientific effort to collect, review, categorize and synthesize existing and relevant scientific research result on evaluation, and it is a relatively recent publication, it can be regarded as a suitable starting point of further exploration.

The first approach enlisted is the *resource allocation approach*, which is concerned with the efficient use of resources, both prospectively in terms of planning and retrospectively in terms of how resources have been used. Researchers and practitioners of this approach claim that a systematic estimation of intervention costs is key to intervention evaluation (Johns et al, 2003), and that it is the accountable and efficient use of resources that provide grounds for evidence-based decision making as a component of good management practice (Smith et al, 2012). In my view, this approach may provide judgement on a simplified input-output model, which have certain limitations in terms of what the intervention have achieved on the ground, however, with strong view on efficiency and future decision making.

The second approach is the *standards or target based approach*, which is concerned to measure success and performance by the application of criteria. Using this approach assumes that one overall target can be identified which is adequately broken down to lower levels enabling evaluation judgements in management interventions (Kessler, 2003). In manufacturing, for instance, drawing up the causal chain of interdependencies between target variables allows for a faster and more objective opportunity to evaluate the effects and significance of disturbances (Knüppel, 2014). This view is often cited as a means to carry out evaluation on effectiveness, but with less emphasis on how efficiently multi-level targets could have been achieved.

Thirdly, the *explanatory approach*, which is concerned with explaining programme impacts and success and make causal statements about what works, when and how. It is a rather rare and difficult task to combine the traditionally diverging research structures of evaluation and explanatory analysis (Rubin, Babbie 2010). In my view, this approach is much more complex than previous ones, bearing strong marks of learning intentions, concentrating on impact routes, mechanisms and causality, however, this complexity is seldom expected.

Fourth, the *formative or change oriented approach*, which provides positive and more complex feedback to support monitoring self-correction during the life of a programme

(George, Cowan, 1999). I strongly support this view personally, in agreement with the vast majority of evaluators, as formative evaluation have the power to underpin change intentions and directly induce alterations as opposed to summative evaluations. In my personal view, evaluation is about learning, and knowledge gathered has to be put to use which may lead to changes in interventions, or the systems they are operated in. I will cover formative evaluation in more depth in later Sections.

And last, the *participatory/development approach*, which seeks to develop networks, communities and territories through bottom-up, participatory methods. "Facilitation skills are essential to ensuring a good quality process, which in turn may require additional resources for building capacity." (Gujit, 2014). In my view, such evaluations (cross-border cooperation evaluations and large territorial cohesion evaluations are good examples) are actually means of bridging cultural gaps and building rapport through establishing a common and compulsory platform to debate, confront and share ideas and form a mutual future. It determinately applies a range of soft elements and techniques, such as story-telling or social mapping, focus groups, facilitated workshops and coaching practices. (Gujit and Gaventa, 1998). In such evaluation cases, evaluation is justifiably a means and not an end (Patton, 1997). These may be positive examples of evaluations conducted to support hidden agendas for greater good.

Though evaluations might use any of, or a combination of the above approaches, they most often stick to a primary approach focus, which – in the EU context – has a *formative* character.

According the Potter (Potter, C., 2006), it is the philosophic approach that influences the choice of evaluation methodologies, for which he defines three categories: the *positivist*, where the programme aspects are „objective, observable and measureable” which require an evidence-based approach; the *interpretive*, where the „evaluator develops an understanding of the perspective, experiences and expectations of all stakeholders” and finally the *critical-emancipatory*, where there is a “participatory, yet critical” approach to the programme, considering the broader context and use of the programme. This classification is rather general and have been further detailed and criticised by many sources. Critics of this classification claim that the critical-emancipatory paradigm can be distorted by ideological or political approaches (Weiss, 1999), while (Louw, 1999) claim that this paradigm does not consider issues of capacity limitation and prioritisation.

## **Professional organisations**

Supranational organisations involved in aiding activities (e.g. OECD DAC, United Nations Evaluation Resource Center and Evaluation Group) have developed their own approach and definition to evaluations.

As per their counterparts and also supporters, evaluators have founded international societies that act as think-tanks, professional workshops and quality standard setting organisations for evaluation activities. The two most prominent on a continental scale being the European Evaluation Society (EES) and the American Evaluation Association (AEA). The approach of these international organisations, and their definition and scope to evaluation provided the foundations for the European Union definition and scope to evaluation approach.

The earlier referred (see Section 2.2.3) *OECD Development Assistance Committee (DAC)* issued their Principles for Evaluation of Development Assistance (OECD, 1991) that remained in power since 1991. According to OECD “An evaluation is an assessment, as systematic and objective as possible, of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, developmental efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors.” (OECD, 2008)

The *UNDP Evaluation Group (UNDP EG)* developed the “Handbook on Planning, Monitoring and Evaluating for Development Results” (UNDP, 2009) which sets high level norms to evaluation.

The *American Evaluation Association (AEA)* has also created their set of Guiding Principles (AEA, 2004) for evaluators. The order of these principles does not imply priority among them, priority varies by situation and evaluator role. Compared to the OECD and UNDP EG guidance, it has a more practical view, though much more general than the others.

The next table summarises the main principles that an evaluation has to pursue according to the above-mentioned organisations.

**Figure 7 – Evaluation principles by determinant evaluation communities and their summary**

<b>AEA</b>	<b>OECD</b>	<b>UNDP EG</b>	<b>Common grounds</b>
Systematic Inquiry Competence Integrity/Honesty Respect for People Responsibilities for General and Public Welfare	Open Results made widely available. Used In partnership Impartial and independent With expertise and independence Transparent Relevant and useful Presented in a clear and concise way. Timely	Independent Intentional Transparent Ethical Impartial Of high quality Timely Used	<b>Objective</b> <b>On purpose</b> <b>Simple and robust</b> <b>Professional</b> <b>In adequate manner</b>

Source: Own construction and with own summary column, based on AEA, OECD, UNDP EG

It is my understanding, that there are similar elements in the above columns, so that certain clusters may be made up to create common grounds (see Figure 7 above). Integrity and honesty (AEA), and being impartial and independent (OECD, UNDP) and transparent (UNDP) requires evaluation to be *objective*. Objectivity should cover both the evaluators and the method they use. Objective evaluations in my view are critical to avoid distortions so that the evaluation may come to real conclusions and recommendations. For the notion of objectivity, I integrate and use the notions „factual” and „unbiased” (Scriven, 1972) and „that disregards expertise of biased persons” (Scriven, 1997) and „quantitative” (Terstiege, 2012). I do not consider „distance” as I agree with (Scriven 1972) that „Distance does not guarantee objectivity”. According to the summary of the above principles, evaluations should fulfil their pre-set *purpose*, following a systematic inquiry (AEA), they should be relevant, useful and transparent, (OECD) and also intentional and transparent (UNDP). Sharing results of the evaluations has to be easy to understand, *simple and robust* – they should be presented in a clear and concise way so that results could be widely available (OECD), to promote that the evaluations are used (OECD, UNDP) and will serve public welfare (AEA). Evaluation should be on a high *professional* level also assuming flexibility – to be carried out with competence (AEA), with expertise (OECD) and should be of high quality (UNDP). It is also important that evaluations are executed in an *adequate manner*, i.e. respecting people (AEA), being open and acting in partnership (OECD) and in an ethical way (UNDP). To sum it up, what all these approaches have in common is showing high standards of integrity, credibility,

relevance and ethical behaviour complemented by a sharp view on quality. As I agree that evaluation is a key means of further development (Tavistock, 2003), it is important to set high standards for both the evaluator and the commissioner side to reduce risks of loss of credibility or misinterpretation.

### 2.3.3 Common evaluation principles

It is also important to break down the above identified set of common principles from the level of evaluation to the actual *persons* carrying out the evaluations, and the *methods* used in the course of the evaluation. As I chose evaluation methods as the central element of my dissertation, I will use the identified and summarised evaluation principles as a major building block to formulating my hypotheses, focusing on the level of methods, acting as implementing agents, means and vehicles of evaluations.

In order to identify the principles to use for the review of evaluation methods, I compared the summary of principles deriving from professional sources, scientific literature and my own personal work experience. Eventually I could identify four themes to use for further scientific research in the field of evaluation.

**Figure 8 – Evaluation main principles by determinant evaluation communities**

Scientific literature summary	Professional organisations summary	Themes for further research
Resource allocation	Objective	Objective and following purpose
Target based	On purpose	Simple and robust answers
Explanatory	Simple and robust	Well founded and flexible to use
Formative	Professional	In partnership
Participatory	In adequate manner	

*Source: Own construction, based on AEA, OECD, UNDP EG public information*

According to professional organisation aggregated principles, an evaluation has to be objective and on purpose, while scientific sources also back up o that evaluations should be target based and formative. In my interpretation, these two themes can be merged into one, so that *objectivity and the alignment to a given purpose* in an evaluation and its methods is a valid field of further research. Please note that purpose in case of EU mid-term programme evaluations is clearly set in compulsory methodology documents (to be covered in detail in Section 3.3.3).

The relevance of researching *simplicity and robustness* in evaluation answers is backed up by both professional (simple and robust) and scientific (explanatory) sources.

Evaluation have to rely on professionalism and competence, i.e. possessing the right resources (both personnel, tools and techniques) to carry out evaluations. Though not explicitly mentioned, I want to highlight *flexibility* out of the above framework of *well-foundedness* as a sign of real professionalism and profound knowledge. Flexibility of experts and methods means that there is competence in the evaluation for a swift alignment to various and expectedly or unexpectedly changing conditions.

Last, evaluation is not an isolated task to carry out: it has to be carried out using the experience of those within the system and for the sake of its stakeholders. I fully agree that evaluations should definitely be executed *in partnership*, in an ethical and correct manner.

## **2.4 Programme evaluation in EU Cohesion Policy**

EU Cohesion Policy and its programmes represent an ideal basis for evaluation activities, being unmatched funding programmes in terms of both complexity, volume and diversity. Evaluation of Cohesion Policy programmes has been compulsory in the participating member states since 1988.

### **2.4.1 EU programme evaluation**

Evaluation is a support activity of the management cycle of EU programmes. Evaluation is „Judgement on the value of a public intervention with reference to criteria and explicit standards (e.g. its relevance, its efficiency). The judgement primarily concerns the needs which have to be met by the intervention, and the effects produced by it. The evaluation is based on information which is specially collected and interpreted to produce the judgement“ (MEANS, 2000). This definition is followed in the relevant EU regulations (EC, 1998; EC 1999a-g; EC2006g-k), the corresponding methodological Working Papers and Working Documents (EC, 2000a-j, 2006a-f) and is generally in line with the OECD definition (OECD, 2002).

According to MEANS, which is a standard sourcebook of evaluation activities carried out in relation to EU programmes, evaluation can be formative or summative. The methods

analysed in this dissertation definitely address and support formative evaluation, as part of the later detailed mid-term evaluation process. Using the widely acknowledged MEANS definition, formative evaluation is: „Evaluation which is intended for managers and direct protagonists, in order to help them improve their action (feedback). Formative evaluation applies mainly to a public intervention during its implementation (on-going or intermediate evaluation). It focuses essentially on implementation procedures and their effectiveness and relevance.“ (MEANS, 2000). On the other hand, summative evaluation uses a stock, and not a dynamic approach (with evaluation questions such as “What has been achieved?”) and focuses more on accountability, compliance and accrues corrective actions. According to a widely acknowledged definition “It is conducted after completion and for the benefit of some external audience or decision-maker (e.g. funding agency, historian, or future possible users). For reasons of credibility it is much more likely to involve external evaluators than is a formative evaluation”. (Scriven, 1991)

Formative evaluation concentrates on the learning aspect of the activity, with questions such as “How can we improve programme performance?”. Evaluation is normally both summative and formative and the balance between the two may well shift during the evaluation (Tavistock, 2003). The distinction between formative and summative evaluation has wittily been summarised in the following way: "When the cook tastes the soup, it's formative; when the guests taste the soup, it's summative" (MEANS, 2000). Please note that both summative and formative evaluations are relevant and useful for public sector interventions, the choice on the combination of approach is determined by the purpose and context of the evaluation.

#### **2.4.2 Evolution of evaluation in the European Union**

Since the reform of Structural Funds in 1988, evaluation activity became a focus area of methodological development as evaluation is regarded as the main means of learning programme mechanisms. As a result, all member states were encouraged to build up evaluation capacities, and develop an approach that bridges the gap between compliance and early impact assessment and focuses on meeting programme objectives thus contributing to Community welfare. Taking on evaluation activity as an integral part of the policy cycle started in the 1989-1993 period and was reinforced in terms of both requirements and methodology in the 1994-1999 programming era.

In the 1989-93 period, ex ante and ex post evaluation were set as requirement for member states (Tavistock, 2003). Member states responded differently to this requirement, and as a result, competence and capacity in the fields of monitoring, evaluation and impact assessments were in a seed phase and not harmonised at all. Based on the quality of the evaluations undertaken in this period, the Commission (in partnership with the European Parliament and the European Court of Auditors) took action to reinforce and converge evaluation capacities in the Community.

As a result, in the 1994-1999 period, ex ante and ex post evaluations were complemented by the introduction of the compulsory mid-term evaluation, with a twofold objective: to act as a basis for programme improvement in the given period and to draw conclusion for the next. The focus of programmes moved to the direction of quantification, clearer intervention logic and accountability, which provided an ideal environment for the development of evaluation. As part of the activities the Commission intended to show an example by establishing Evaluation Units within the organisation of all Directorate General (DG) to act as coordinators and provide guidance for evaluation activities in member states.

The Commission also started the MEANS programme in 1999 with the intention to collect and synthesise best practices in evaluation in all over Europe. By introducing a new system of requiring member states to undertake ex ante evaluations by independent evaluators commissioned by the Commission, the EC could have an independent quality assurance support in the official negotiation process which proved to be valuable. The MEANS project was the first and most notable milestone in the development of evaluation in Europe. By providing evaluators, commissioners, member state officials with a general approach and a toolbox that was in line with the relevant regulation and the expectations of the Commission, a new era in evaluation has started.

In the 2000-2006 period, the next reform of the Structural Funds introduced new approaches. Member states were required to initiate and manage their evaluations. New focuses appeared that were intended to promote the principles of environment protection, labour market and gender equality. The introduction of the institution of performance reserve created another means to favour interventions with good performance. Within this period started the second era of evaluation, by issuing *The Guide* (Tavistock, 2003) by the Tavistock Institute that continued, fine-tuned and further developed the approach set

in the MEANS volumes according to the practical experience and scientific achievements gained in the meantime. In line with the intention of the EC to provide continuous methodological assistance and guidance to evaluators, Working Papers (EC, 2000a-j) and Working Documents (EC, 2006a-f) were designed.

In the 2007-13 period, the Commission introduced no major changes in terms of the ex ante and mid-term evaluations. As a slight modification to ex post evaluations, however, it required the member states to carry out the assessment in the year of the programme close (as a result of the so-called n+2 rule<sup>3</sup>, the 2007-13 period is practically a 2007-15 period in view of potential implementation activity). In line with the intention of the EC to provide continuous methodological assistance and guidance to evaluators, and as a preparation for the 2007-13 period, Working Documents (as opposed to the Working Papers in the 2000-2006 period) were designed in 2006.

As for the currently ongoing 2014-20 period, regulations governing evaluation activities hardly changed. It is important to note however, that planning frameworks became very strictly centrally driven to follow the 11 thematic objectives and investment priorities of the funds all representing EU2020 objectives (EC, 2014a-b). As a result, the frameworks of planning are more strict but more targeted to real EU level needs than ever. Evaluation is expected to be more rigid, much more to the point, focusing on the clarity and justifiability of intervention logic and quantification of objectives. The performance reserve is also strict, stating that 6% of funds allocated are at risk if the progress of the achievement of programme objectives is lagging behind on time proportional basis, no matter if it was caused by planning, implementation or other errors. However, this amount is welcome to be spent on top-performing interventions.

The next table provides a simplified and focused summary of the one that can be found in the 2013 edition of the Guide (Tavistock, 2003), providing a brief overview on the relevant changes in evaluation requirements through EU budgetary periods.

<sup>3</sup> According to EC definition, n+2 and n+3 rules relate to financing rules for the annual allocation of money from the European Union's Structural and Cohesion Funds. If the funding in question has not been spent by that date, the Commission can decommit future budget allocations. Automatic decommitments are made if funding is not spent, or requests for payments are not made, by the end of the second year (n+2). This deadline was extended to three years (n+3) for the EU-10 member states plus Greece and Portugal until 2010.

**Figure 9 – Development of evaluation activities over the programming periods from 1989 to 2020**

Period	1989-1993	1994 - 1999	2000 - 2006	2007-2013	2014-2020
<b>Ex ante</b>	Required from member state	Required in partnership (Member state and EC)  Must include environmental assessment	Required primarily from member state.  Special attention to environment, labour market, gender equality	Required primarily from member state.  Special attention to environment, labour market, gender equality	Required primarily from member state.  Special attention to environment, labour market, gender equality  Strong focus on intervention logic
<b>Mid-term</b>	Not required	Required from programmes with 3+ years of duration.  Focus monitoring data and objective achievement	Required from MA in partnership with EC.  Evaluator must be independent  Evaluation in 2003 and 2005.	Required from Member state.  Focus shift from compliance to needs based approach.	Required by priority  Focus contribution to objectives  Report required by 2021.
<b>Ex post</b>	Required on national level	Required in partnership with EC  Focus: impacts	EC responsible  Independent evaluator  Within 3 years from programme close	EC responsible  Independent evaluator  Due in 2015	EC responsible  Independent evaluator  Due in 2023
<b>Performance reserve</b>	na	na	By 31 March 2004, 4% held back	Optional for member state	6% of allocation

*Source: own edition, based on the Guide (Tavistock, 2003)*

*Legend: MA: Managing Authority, EC: European Commission*

### 2.4.3 Programme evaluation subjects: interventions

There are many theoretical and practical approaches to the evaluation of intervention programmes. The development of EU evaluation guidance documents (i.e. Working Documents, Working Papers, the MEANS), took account of many such approaches. Out of the many available I want to highlight the ESRI approach that has been a general approach in Ireland and has made its way to become an important part of the Working Document No.1 (EC, 2006a) of the EU Commission. It has been widely used in EU co-

financed programme evaluations in Central and Eastern Europe and other state intervention programmes and often regarded as a cornerstone to policy and intervention evaluation.

The ESRI approach (ESRI, 1993) was originally developed as part of the mid-term evaluation of the CSF for Ireland 1994 to 1999, carried out by the Economic and Social Research Institute (ESRI) in 1997. Subsequently, it was used in a number of other CSF-level evaluations in Ireland, including in the ex ante and mid-term evaluations of the NDP/CSF 2000-2006. The methodology is emphasized in the European Commission's working paper on ex ante evaluation (EC, 2000a) as a useful approach for the evaluation of any public intervention.

The theoretical basis for the methodology rests on the widely accepted economic principle that public investment should be directed towards addressing distortions or market failures in the economy. In the words of the UK Treasury, "a necessary but not sufficient condition for government intervention to improve economic efficiency is that there is some form of market failure". If the market can do it, there is no need for the Community to intervene. The approach provides a basis for classifying intervention programmes and their components on the basis of identifying the nature of the market failure or distortion that they seek to address.

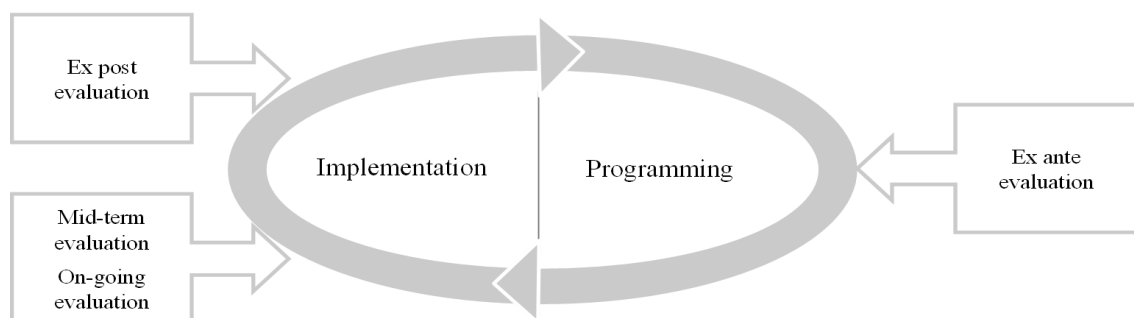
The methodology employs a four-fold typology of measures (potential subjects to policy and programme evaluations) corresponding to the following instances of market failure. *Public goods* are those that are "non-rival" in consumption. One person's consumption of the good or service in question (e.g., a public park, street lighting or a road in a rural area) does not preclude its consumption or use by others. Because of the difficulty or impossibility of charging for the use of such goods, the market will tend to under-provide or, in some cases, not to provide them at all. A *corrective subsidy* is an open-ended intervention designed to alter relative prices facing firms or individuals in order to correct for, or capture, a general or ongoing externality or spill over effect. For example, firms will tend to under-invest in the provision of general (i.e., not firm- or job-specific) training for their employees because some of the benefits cannot be retained by them but rather accrue to the trainees or to other firms (in the event that the trained employees leave). A *targeted subsidy* is designed to alter behaviour where private agents are thought to be inadequately informed, or where a specific externality or information barrier exists (for

example, firms may not recognise the benefits deriving from investment in research and development or management training). Unlike the second type, these are not open-ended subsidies. And last, *redistribution* covers interventions where the primary motivation is redistributive, i.e., to provide goods or services to groups or regions which would otherwise not have access to these. The allocation of interventions to the various categories is merely the starting point for a programme (or underlying policy) evaluation. Once allocated, the measures are examined under a series of evaluation criteria or questions designed to test the robustness of the preliminary market failure justification. Thus, the fact that a particular measure is classified under one of the categories does not, in itself, justify investment in the area concerned.

#### 2.4.4 Programming, implementation and evaluation cycle

Evaluation of EU-funded programmes is mandatory “Evaluations shall aim to improve the quality, effectiveness and consistency of the assistance from the Funds...” (EC, 2006g). Programming, implementation and evaluation together make up an integrated cycle. Programming uses the inputs deriving from evaluation (lessons learnt on previous intervention), then implementation starts, whose results will be assessed through a next evaluation to further enhance the next round of programming. As part of the evaluation cycle closely linked to the programming cycle the standard EU programme evaluation activities are *ex ante* evaluations (used in the programming and planning phase, *mid-term evaluation* / *on-going* evaluations (used during programme evaluation) and *ex post* evaluations (used at programme close). (See Figure 5 and 10 for details). For easy reference on how programming, implementation and evaluation phases are interlinked, please refer to the figure below:

**Figure 10 – Programme and evaluation cycle of the Structural Instruments co-financed programmes**



Source: own construction

All three evaluation phases (i.e. ex ante, mid-term and ex post) and their tight interlinkages promote the relevance and effectiveness, efficiency, and accountability in programme implementation.

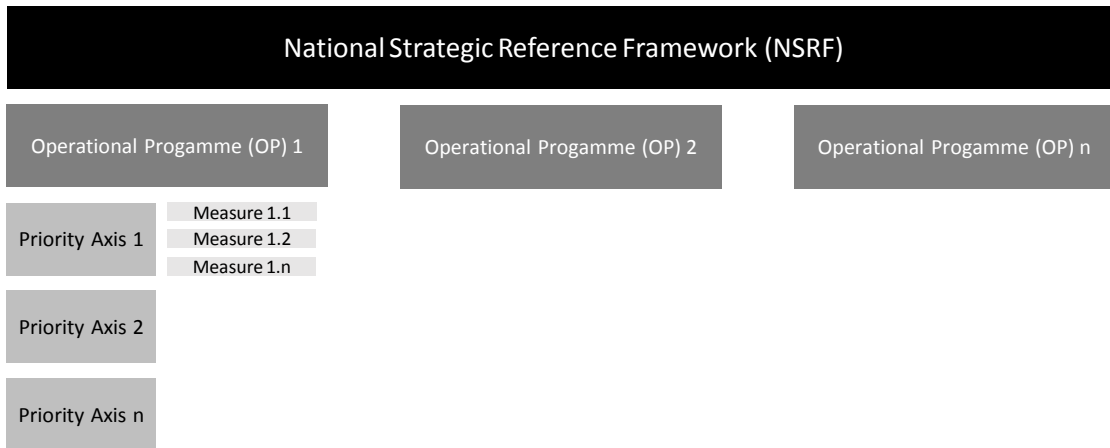
The driving factor for *the choice of mid-term evaluation* phase methods in my research scope lies in three points: First, mid-term evaluations carry strong development policy aspect, and within that a sharp focus on the institutional aspect, i.e. relevance and rationale, efficiency and effectiveness. Second, mid-term evaluations cover the most areas by assessing the performance of the actual period retrospectively while they also formulate recommendations for the next programming period. Third, 2007-13 mid-term evaluations are the latest relevant (there are 2014-20 ex ante evaluations already available at present, but their goal, scope and therefore their choice of method is different) approved evaluations available. On the other hand, 2007-13 ex post evaluations are just to be commenced in member states, however, they are concentrating on the financial and administrative aspects of the development programmes in a summative view, not the institutional aspects with a formative approach.

#### **2.4.5 Structure of Cohesion Policy interventions**

In the 2007-13 programming period each member state developed its Community Support Framework (CSF), a National Strategic Reference Framework (NSRF) or in case of member states with total funds budget below EUR 1 billion, a Single Programming Document (SPD) (hereinafter both referred to as NSRF) to act as country level plan for the EU budgetary period. The NSRFs took over the role of National Development Plans that served similar purpose for the previous, 2000-2006 period. (Please note, that the required content of the respective plans had undergone a transformation, though main elements remained very similar.)

The NSRF articulated country level objectives, means and budgetary allocations for the period. They drew their inspiration from the priorities adopted by the member states in October 2006 in the “Community Strategic Guidelines for Cohesion” (EC, 2007). The NSRF created a consistent structure of interventions, breaking down top level objectives to Operational Programmes (OP), Priority Axes (PA), and lower level interventions (measures, projects). They contain strategies in detail, identifying the investment plans, the regions targeted and the impacts expected (EC, 2007).

**Figure 11 – NSRF-OP-PA-measure**



*Source: own construction*

Please also note, that the so-called SPDs have been used by countries with a total programme budget not exceeding EUR 1 billion. SPDs – similarly to NSRFs, but in a more concise and simpler form – were required to include the description of the current situation, strategy and priorities, an integration of the ex ante evaluation, an outline of the measures, an indicative financial plan, the ex ante verification of additionally, partnership considerations and provisions for implementation.

It is important that the NSRF should be based on available national regulation, policy documents and strategic documents. The NSRF should be very clear in defining which areas require Community intervention and support, and which areas are to be handled via national programmes and resources in order to fulfil the requirement of additionality.

However, this has not been the case in programmes later detailed in my research, as NSRFs and OPs in Bulgaria, Romania and Malta have only partly been built on existing national strategies along with the majority of the EU-12 countries. The main reason was lack of approved national strategies and the lack of awareness of the EU intervention logic, as a result of no relevant experience. In these countries the building of the NSRF and therefore regional policy rather followed the opportunities created by Community funding, resulting in programmes with less strategic focus, but with a larger potential to absorb funding. This is due to the fact that budgetary constraints in these countries resulted in a dramatic drop in available financial resources for development purposes, and therefore it was exclusively driven by EU funding.

#### **2.4.6 Cohesion Policy programme evaluations in the 2007-13 period**

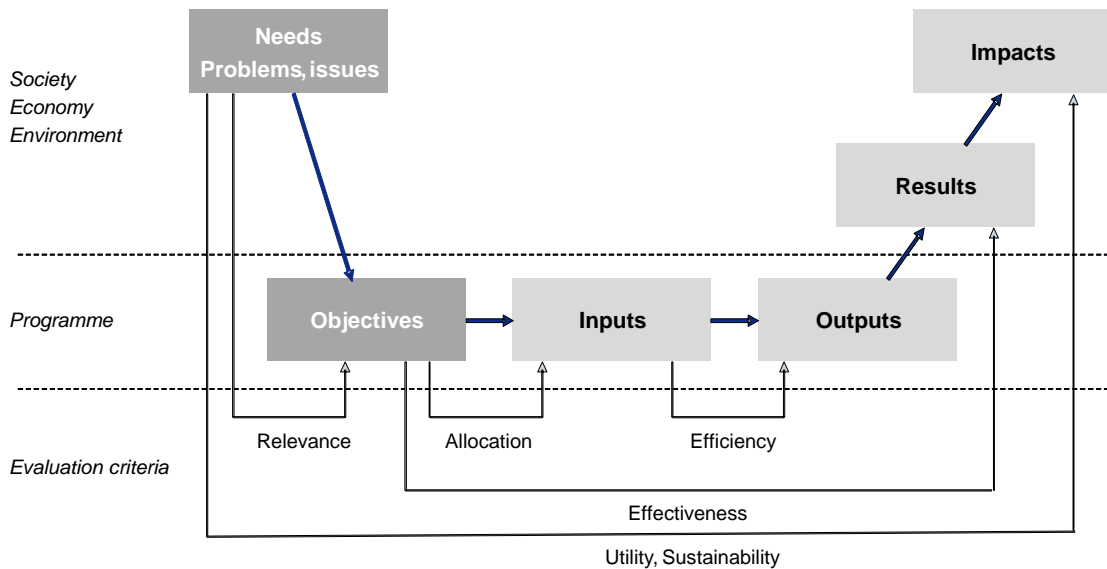
Evaluations in the 2007-13 period have been widely used by the European Commission to learn more about intervention mechanisms in an enlarged Europe, where not only 15, but the considerably different newcomers, the EU-10 countries have just accomplished their first half programming period (i.e. they all had 2004-06 NSRFs and OPs in the 2000-2006 programming period as a result of joining the EU on 1 May 2004), where their new experience provided new grounds and new aspects for evaluation, left alone Romania and Bulgaria joining in 2007 with EU funding experience limited to pre-accession programmes (IPA). With this opening perspective, evaluation gained more focus than it previously had. Evaluations were aimed at various subjects of the support system, e.g. policies,, programmes (NSRFs and Operational Programmes and their respective priorities, measures, interventions), projects have been assessed. Evaluations concentrating on horizontal issues (which has gained more emphasis in the 2007-13 period, in line with Community requirement to focus more on environment, labour market and gender equality), thematic evaluation (e.g. SMEs, disadvantaged groups, or cross-cutting themes, like IT or RTDI) have risen in number. Apart from the programme aspect, more evaluations concentrated on the implementation aspect, i.e. the institutional system (from planning to execution, controlling, monitoring, reporting and audit), and processes (efficiency, lead times, capacities, costs, bottlenecks and built-in controls).

Evaluation topics varied greatly in terms of sector (economic development, transport, environment, capacity building, regional development, agricultural / rural development) and timing, in line with EC obligations (ex ante, ongoing / mid-term, ex post) and other (so-called ad-hoc evaluations). The main evaluation criteria assessed remained practically untouched compared to 2000-2006 period, with concentration on rationale, relevance, effectiveness, efficiency and impact/ sustainability (EC, 2006a-f).

#### **2.4.7 Cohesion Policy evaluation methodology framework in the 2007-13 period**

Following the considerations of the evaluation framework presented in European Commission's '2007-2013 Working Document No.1.: Indicative guidelines on evaluation methods: Ex ante evaluation' (EC, 2006a) is a solid reference point for presenting Community approach to programme evaluations, being a process-based framework:

**Figure 12 – European Commission evaluation framework 2007-13**



Source: EC, 2006, own edition See footnote for definition of labels<sup>4</sup>

The main elements of the above evaluation framework constituted a Community-wide nomenclature for evaluation notions. The process encompasses from the needs, problems and issues that are translated into objectives that require inputs to address, creating immediate outputs, longer term results and finally, impacts. Evaluation concentrates on the comparison between relevant stages along the programme planning and implementation process.

4

**Needs:** an environmental, economic or social gap initiating a case for intervention

**Objectives:** hierarchy of objectives with the aim of fulfilling environmental, economic and social needs

**Inputs:** physical, human and financial resources allocated for the achievement of the objectives

**Outputs:** the direct consequences of the utilisation of inputs during the achievement of objectives

**Results:** the immediate consequences of the outputs affecting primarily the direct stakeholders

**Impacts:** the long-term consequences of the outputs affecting direct and indirect stakeholders as well. Expectedly, the impacts should succeed in fulfilling the needs.

The evaluation criteria examine the link between the elements of this process, such as

**Relevance of the objectives:** whether the objectives reflect to real needs that are significant and are in strategic focus at community or at national level

**Budget allocation:** whether the allocation of the resources at disposal (community and national public financing) is optimal

**Efficiency:** whether the transformation of the inputs into outputs is optimal

**Effectiveness:** whether the results and impacts fulfil the goals set and the original needs

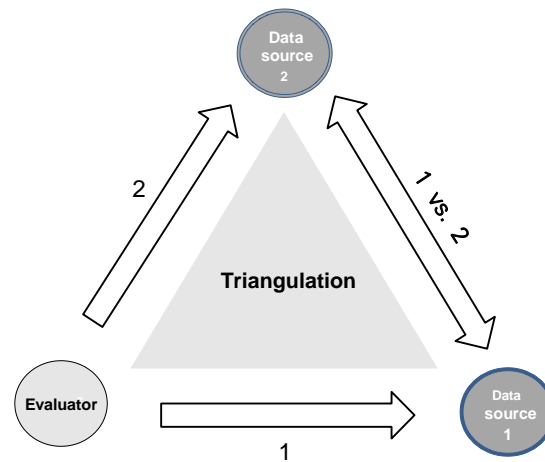
**Utility:** whether the intervention has a net utility for the local society, the region, the country or the European Union

**Sustainability:** whether the outputs, results and impacts of the intervention are sustainable in economic, environmental and social terms.

## Triangulation approach

The triangulation concept is a general EU approach to evaluation activities, in which findings and conclusions have to be justified using at least two sources of information in order to establish a potentially adequate level of reliability and robustness. Triangulation in practice usually means that the evaluator uses multiple data sources in the course of the evaluation and relies on findings that are backed up by the consent of the information sources.

Figure 13 – Triangulation concept



Source: own construction

Evaluators tend to follow an evidence-based (i.e. based on facts and figures) and an opinion based approach (i.e. subjective sources, like interviews, workshops) in order to triangulate the findings and conclusions where applicable. All three methodologies selected for my research provide an evidence-based approach, therefore, provide a sound basis for further evaluation activity. Setting the scenes for evaluation using fact-based information provides both excellent grounds for an objective assessment and the potential for narrowing (focusing) the scope of evaluation for reasons of cost-effectiveness (e.g. more targeted consultations and questions, pre-set topics to explore, focusing capacities with a high potential to disclose evaluation answers).

## 2.5 Mid-term Cohesion Policy programme evaluations in 2007-13

Mid-term evaluations and on-going evaluations are often quoted as synonyms (EC, 2006a), and in this study the notion of mid-term evaluation also covers on-going evaluations. Mid-term evaluations are evaluations carried out in the course of programme implementation (usually in mid-time or in the second half of programme duration), aiming at identifying potential interventions that can be promptly used in order to enhance the relevance, effectiveness and efficiency, accountability of the programme, and collecting substantial information to the formulation of the next programming period intervention plan.

### 2.5.1 Evaluation context

The definition, context and background of evaluations have already been discussed in Sections 2.1-2.3. However, it is important to describe, how a methodology fits into the programme evaluation activity. The objective of programme evaluations is to answer Evaluation Questions (EQs). Such EQs form part of the Terms of Reference (ToR) which sets out all relevant information about the evaluation project. The EQs address the evaluation themes that are relevant to the actual implementation stage of the programme (i.e. ex ante, mid-term / on-going or ex post) and those other issues that the institutional system representatives or other stakeholders of the evaluations are particularly interested in to explore further.

### 2.5.2 Evaluation themes

As previously described, mid-term evaluations use a series of methods, tools and techniques to draw meaningful conclusions to answer the predetermined evaluation questions. Mid-term evaluations in general are concentrating on the following evaluation themes: relevance, effectiveness and efficiency and accountability. Each of these evaluation themes can be tackled through a wide range of tools (*I used the example of road constructions to facilitate understanding in the following sub-sections*).

#### **Relevance**

The evaluation of relevance concentrates on the rationale of the programme in view of the current and foreseen external environment and programme context. This is a focus area in ex ante evaluations, but also a very meaningful issue in mid-term evaluation, to gain a retrospective justification on the relevance of the then planned interventions. There are usually more issues in the scope of the relevance evaluation about programme objectives and corresponding selection of indicators, programme scope and the quality of past predictions.

However, in the mid-term phase, the focus area is continuous relevance, aiming at the exploration whether previously set programme objective may remain valid and untouched if changes in the external environment have not reached a significant level (*e.g. major changes in the demographic concentration or the establishment of new economical centres might require modifications of the foreseen road construction map*).

Internal consistency of programme objective hierarchy and external coherence with higher level relevant community or national policies and regulations are assessed under relevance as well.

Evaluation methods corresponding to the above issues encompass for instance

- *Consistency assessments* (internal and external) regarding the environment, structure of objectives (if the breakdown of objectives still favourably follows the MECE principle, i.e. mutually exclusive and collectively exhaustive).
- *Indicator quality analysis*, checking if indicators are capable of translating objectives to measurable performance units following the SMART, QQTTP or other criteria set.
- *Selection criteria analysis*, examining if the project selection parameters are capable of reflecting higher level objectives.

### **Effectiveness**

The evaluation of effectiveness concentrates on programme results, which has at least two aspects: results on the ground (in natural units, *e.g. kms of roads constructed*) and financial results (in currency, *e.g. fund committed, contracted or spent on road construction*). Standard evaluation issues include real progress, financial progress and expected impacts. In a few cases, consistency of objectives is included in this evaluation theme also as this topic is the “vehicle” of the theoretical infrastructure creating the opportunity for effective programme implementation.

Evaluation methods used for measuring effectiveness are for instance:

- *Simple Progress Overview Tool*, a tool for investigating financial progress of a programme and revealing potential bottlenecks with a concentration on absorption and related ratios e.g. if there is an administrative shortage in any stages of the application process resulting in low absorption figures or ratios, see Section 3.4 for details.
- *Multi-level mapping*, a method for mapping all relevant facets of actual programme status in terms of natural progress, e.g. regional split, industry split, beneficiary split, project volume split.
- *Prognosis of financial progress*, using trend analysis and mathematical extrapolation on the basis of defining previous trends in absorption and

prolonging this trend coupled with considerations on n+2 and n+3 rules and budgets available.

- *n+2 / n+3 compliance assessment*, focusing on the scheduling of fund allocation so that the n+2/n+3 rules are not violated resulting in potential decommitment of funding in certain years of the implementation period.

### **Efficiency**

The evaluation of efficiency concentrates on the assessment of the relative yield of the implementation (output/input) as a result of the operation of the institutional system. Efficiency in EU programmes have at least two relevant aspects, one being the relative cost of producing one unit of the respective indicator (*e.g. cost of producing one km of road*) the other being the cost of fund allocation / fund (*e.g. cost of deploying one EUR of funding for road construction*)

Efficiency assessment might use a wide variety of methods, for instance

- *Lead time analysis*, a tool for measuring relevant time durations in the implementation process and identifying the process factors that play a role in creating bottlenecks in the system (see Section 3.6 for details)
- *Capacity, capability and cost assessment*, a tool for identifying input factors of the implementation system and making internal and external comparisons.
- *Project quality review*, a tool for assessing the errors in the applications pointing at either competence issues on the side of potential beneficiaries, or lack of clear and simple documentation provided by the institutional system.

### **Accountability**

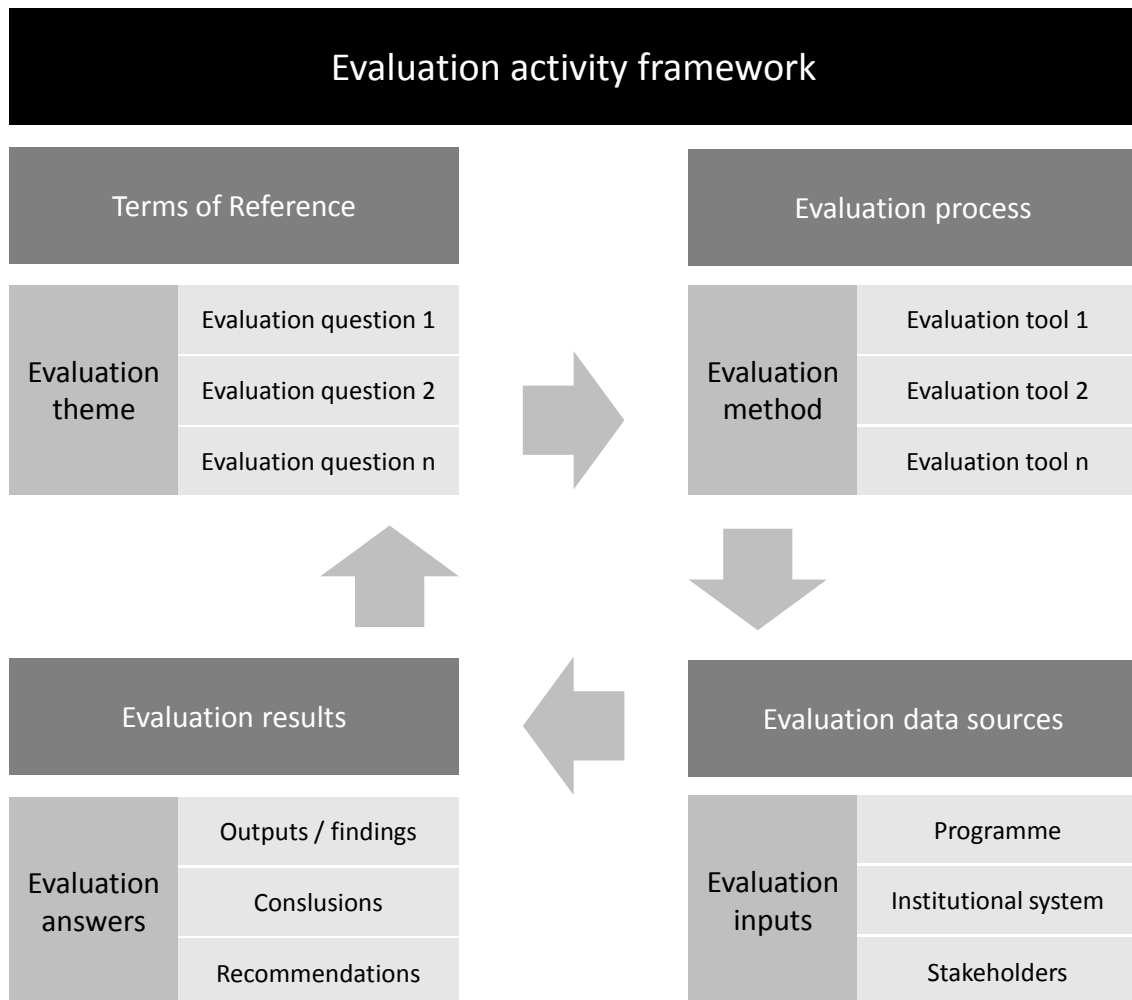
The evaluation of accountability focuses on the extent to which implementation follows the considerations laid down in the respective Community and national regulations. The evaluation of accountability may rely on methods such as

- *Post Contractual Assessment*, a tool for identifying which process stage and why is responsible for lag times in the implementation system and what proportion of funding is at risk of decommitment (see Section 3.5 for details)
- *Regulatory compliance assessment*, a comparative analysis of the programming and/or implementation and the respective general regulations.

### 2.5.3 Evaluation methods

Evaluation methods are selected and applied to tackle evaluation themes and questions. Evaluation methods use tools and techniques which, in the end, are expected to provide answers to the evaluation questions. Based on my working experience I drew up Figure 14 on the next page, framing the system in which evaluation activity is conducted. This framework is rather general, and regardless of programme type and country.

**Figure 14 – Evaluation activity framework**



*Source: own construction based on work experience*

The structure of evaluations therefore follows this pattern: the Terms of Reference set out the evaluation themes including evaluation questions („WHAT we need to learn”) that are addressed by evaluation methods („HOW we can learn that”) and tools („WHICH means to use to learn that”). Eventually, it is the output of using the methods and the tools and techniques linked to them that addresses the evaluation question and eventually answers it.

It is important to highlight that evaluation is a one-off, but a recurring activity. Although evaluations set out to answer evaluation questions, the answers tend to yield more questions and shed light to topics that have not been included in the scope of the evaluation. Therefore, it is evaluation results *inter alia*, that provide the grounds for the commencing of the next evaluations, influenced by forthcoming evaluation stages (*ex ante*, mid-term, *ex post*) and corresponding regulations and guidelines.

## **2.6 Formulation of research scope and hypotheses based on the summary of literature review**

The scope of my research is international comparative assessment of selected mid-term institutional evaluation methods and regional programmes in the EU-12. Having covered relevant scientific, professional and EU literature on what factors contribute to better evaluation, I intend to identify and formulate the hypotheses I am testing in my research.

### **2.6.1 Research scope**

The scope of my research is three well-acknowledged mid-term programme evaluation methodologies applied in - *inter alia* - five mid-term evaluations of National Strategic Reference Frameworks (NSRFs) and Operational Programmes (OPs) in 2007-13 in three EU-12 countries: Bulgaria, Malta, and Romania, with an underlying outlook to EU-12 EU co-funded regional development programmes and regional development measures and their results.

*Cohesion policy* programmes (NSRF, OP) are the primary, if not the only development resource in the EU-12 countries. Therefore, they provide excellent testbed and offer the most potential results to research programmes, evaluations and evaluation methods. As the main purpose of Cohesion Policy is decreasing regional development gaps, it is important to evaluate intervention *results*, how progress has been made toward this goal. From this perspective, it is the *regional operational programmes* that explicitly address to enhance regional development and minimize disparities, therefore, they provide the best subjects for my assessment.

The reason for the selection of *mid-term evaluation in 2007-13* is that mid-term evaluations carry strong formative, development policy aspect, and within that a sharp focus on the institutional aspect, i.e. relevance and rationale, efficiency and effectiveness.

Second, mid-term evaluations cover the most areas by assessing the performance of the actual period retrospectively while they also formulate recommendations for the next programming period. Third, 2007-13 mid-term evaluations are the latest relevant (there are 2014-20 ex ante evaluations already available at present, and 2007-13 ex posts are under way, but their goal, scope and therefore their choice of method is different) approved evaluations available.

The selection of *institutional evaluation* is closely linked to the previous explanation. Mid-term evaluations in the 2007-13 period have a strong focus on efficiency (in compliance with the Community guidance (EC, 2006a)), and the evaluation of efficiency essentially assesses the performance of the institutional (management and implementation) system.

*Methods* are both implementing agents and building blocks of evaluations. Adequate and adequately used methods may provide sound evaluations. Therefore I chose to assess methods according to the principles and aspects identified by narrowing and distilling the findings of the scientific literature review and my own experience.

The reason for choosing *EU-12 countries* was practical: the accession of these countries in the EU implied the most changes, also in terms of Cohesion Policy. With Hungary being one of them, I had a special interest to use these countries as a basis for comparison on the application of methods. Also, I had personal work experience with their country context, programmes, evaluations and methods.

### 2.6.2 **Research approach**

It is required but not sufficient – according to both scientific sources, practical feedback (see Section 2.3.3) and my own experience – to conduct Cohesion Policy evaluations that fulfil a purpose, i.e. linked to the assessment of general intervention objectives, and which use methods that are objective, simple and robust, and flexible. Methods, as means of evaluation have to fulfil similar criteria. The adequate use of sound methods considerably contributes to better evaluations. Similarly, regional programmes act as one of the most important means of Cohesion Policy implementation. Well planned and well executed regional programmes contribute to the fulfilment of Cohesion Policy objectives.

Through joining these two logical paths, it is recommended to use a twofold research approach: one direction is represented by methods applied in evaluations, while the other direction is embodied by regional development programmes directly affecting Cohesion Policy results.

### 2.6.3 Research hypotheses

As a result of the concise scientific and practical literature review, coupled with my own international experience, I pursue to challenge five hypotheses (H1-5) within the context of this research. Hypotheses 1-2 cover Cohesion Policy and regional development programmes in EU-12, while Hypotheses 3-5 deal with my main topic, method features.

H1: EU Cohesion Policy could contribute to the development of regions in the EU-12, however, it could not counteract the growing regional disparities within these countries.<sup>5</sup>

H2: Regional Operational Programmes in the EU-12 countries in 2007-13 are diverse in terms of their approach and their relative weight in the NSRF despite their content (interventions) and their mid-term evaluations are very similar.<sup>6</sup>

H3: The selected evaluation methodologies are capable of providing objective<sup>7</sup> basis for evaluation and address at least one of the general intervention objectives of effectiveness (and relevance), efficiency and accountability.<sup>8</sup>

H4: The selected evaluation methodologies provide simple<sup>9</sup> and robust answers to evaluation questions.

H5: The selected evaluation methodologies have the flexibility to be customised to the context of the evaluation.

<sup>5</sup> The negative assumption in the hypothesis is backed up by the Sixth Cohesion Report (EC, 2014d, p.232)

<sup>6</sup> Both diversity (as a result of different economic structure, regional authority system, general approach of central administration to regional development, etc) and similarity (reflected in EC and scientific publications) are assumptions based on previous research and work experience

<sup>7</sup> For the notion of objectivity, I integrate and use the notions „factual” and „unbiased” (Scriven, 1972) and „that disregards expertise of biased persons” (Scriven, 1997) and „quantitative” (Terstiege, 2012). I do not consider „distance” as I agree with (Scriven 1972) that „Distance does not guarantee objectivity”

<sup>8</sup> see MEANS, 2000

<sup>9</sup> I pursue the notion of simplicity in the sense of „easy to understand” and „easy to communicate” to broad public

### **3 INTERNATIONAL COMPARATIVE ASSESSMENT OF SELECTED MID-TERM INSTITUTIONAL EVALUATION METHODS AND REGIONAL PROGRAMMES IN THE EU-12**

#### **3.1 Research set-up for the regional analysis and methodology analysis**

##### **3.1.1 Research overview**

My research is based on five hypotheses to be tested via two international comparative analyses:

1) *regional programme analysis* (Section 3.2 and 3.3).

2) *methodology analysis* (Section 3.4, 3.5 and 3.6), and

For the *regional programme analysis*, Cohesion Policy programmes in general and specifically Regional Operational Programmes (ROPs) in EU-12 countries acted as vehicles to the testing of my first two hypothesis (H1-2) on the results of Cohesion Policy on development and regional disparities in the EU-12, and the diversity of programme approaches and relative weight of ROPs in their NSRFs despite the similar content. The regional programme analysis uses all 2007-13 ROPs in the EU-12, therefore, it does not require a sampling exercise and description to underpin research foundations.

For the *methodology analysis*, I have previously selected three evaluation methodologies that I have developed, tested and applied in many EU countries over the past 10 years. My aim is to conduct an analysis via comparing the experience drawn through the application of these methodologies in diverse country and programme contexts in order to test the last three hypotheses (H3-5) of the research.

In the following sub-sections (3.1.2 and 3.1.3), I will present the country, programme and evaluation selection method for testing the hypotheses in the analyses. Then, I will describe the definition, scope, and usage of the three preselected evaluation methods. In all cases, the method will be demonstrated through at least two real life international examples. Methodology overviews will be followed by an assessment of the evaluation methods concentrating on the testing of the hypotheses, i.e. objectivity and correspondence to evaluation themes, simplicity and robustness, and last, flexibility.

These chapters will be closed by lessons learnt and conclusions of the comparative analyses.

The research set-up (see Figure 15 below) summarises the frameworks of the research in terms of both the methodology analysis and the regional programme analysis and will act as a reference point throughout the research activities.

**Figure 15 – Summary of research set up**

Analysis	Focus	Hypothesis	Research aspect	Subject
<b>International comparative regional programme analysis</b>	All Regional Operational Programmes in the EU-12 in 2007-13	H1-2	Allocation approach Relative weight in NSRF Content (measures) Results analysis	1 All EU-12 regions 2 All EU-12 Cohesion policy programmes 3 All EU-12 ROPs and OPs with regional development measures (altogether 46 OPs + 5 programmes including regional development measures)
<b>International comparative methodology analysis</b>	Three methods (SPOT, PCA, LTA) used in EU-12 mid-term evaluation in 2007-13	H3-5	Objectivity Link to objective Simplicity, robustness Flexibility	1 NSRF mid-term Romania 2 SOP-IEC mid-term Romania 3 OPI mid-term Malta 4 OPII mid-term Malta 5 OPRD mid-term Bulgaria

Source: own construction

### 3.1.2 Selection of research subjects for the methodology analysis

#### Subjects of the methodology analysis: pre-selected evaluation methods

The research concentrates on the presentation, demonstration and international comparative analysis of three evaluation methods and tools I have developed over recent years:

*Simple Progress Overview Tool (SPOT)* is a method and tool for investigating financial progress of a programme and revealing potential bottlenecks with a concentration on absorption and related ratios e.g. if there is an administrative shortage in any stages of the application process resulting in low absorption figures or ratios.

*Post Contractual Assessment (PCA)* is a method and tool for identifying which process stage and why is responsible for lag times in the implementation system and what proportion of funding is at risk of decommitment.

*Lead Time Analysis (LTA)* is a method and tool for measuring relevant time durations in the implementation process and identifying the process factors that play a role in creating bottlenecks in the system.

All three methods are used for the institutional aspect of the evaluation, in line with the title of my dissertation. The table below concisely summarises their focus and the advantages and disadvantages of their application.

**Figure 16 – Summary of methods in the methodology analysis (focus, advantages, disadvantages)**

	<b>Focus</b>	<b>Advantages</b>	<b>Disadvantages</b>
<b>SPOT</b>	<b>Programming and implementation issues behind financial progress</b>	Provides a dashboard on real progress, spotting reasons behind lags.  Clear distinction of “pass” and “process” related issues along the implementation system.	Requires benchmarking data for “pass” ratios, time-proportional values should be interpreted with caution (non-linear tendency)  Limited use with financial instruments and large projects.
<b>PCA</b>	<b>Quantification of components underlying risk probability and risk extent</b>	Provides forecast and detects crucial factors of and potential loss of absorption	Requires good quality (compulsory, i.e. existing) monitoring data in great detail for a sample.
<b>LTA</b>	<b>Institutional reasons behind progress delays</b>	Easy to communicate, quantified and comparable results identifying and quantifying bottlenecks in the processes.	Requires the right choice of benchmark data to interpret adequately

*Source: own construction*

### **Selection approach**

My choice on the selection of the research subjects was considerably influenced by the element to include the evaluations that I personally took part in, so that my personal view and experience can be presented on top of the objective assessment. This aspect has limited the number of countries from twelve to seven, still allowing for a wide range of profiles to research. I used the following three-fold (country, programme, evaluation) selection criteria:

In order to promote *country coverage*, the selected countries are all member states of the European Union. In order to make items more comparable, EU-10 countries (accessing the European Union in 2004) and EU-2 countries (accession in 2007) were both selected, containing larger, smaller and also average-sized countries. For *international comparability* I found it important to demonstrate the same method in at least two separate country cases. As for *country profile* I selected a well-balanced representation of both geographical coverage, diversity (i.e. Central and Eastern European, Southern European Countries) and development.

To ensure *programme coverage*, the selection contains only high level programmes, i.e. NSRFs and OPs, and preferably both for greater variety. For *sectorial coverage*, the

selection contains all types of sectorial interventions: NSRFs carry the opportunity of covering all development areas and topics, therefore they are in the focus. However, as a great proportion of the grants is usually allocated to economic development and regional development, and this is one of the most massively evaluated area, economic and regional development is overrepresented.

*Availability* of items is of utmost importance: the selection only contains evaluations which have been concluded related to the 2007-13 programmes and which are mid-term evaluations in the sense of the Community regulations and which are publicly available. There are only evaluations that are available in English *language* in the selection for convenient processing, avoiding the risk of losing information in interpretation and budgetary constraints of the research. In terms of *scheduling* only 2007-13 mid-term programme evaluations have been selected, as this is the evaluation phase that covers the most evaluation themes, therefore, it has to greatest potential to include elements relevant to the research.

### 3.1.3 Countries, programmes and evaluations selected for methodology analysis

The selection process was an iterative one, following the above-mentioned criteria and parameters. Seven countries (where the methods have been applied) and their respective programmes (NSRFs and OPs) and corresponding evaluations have been assessed. Eventually, the predetermined profile could have been selected, with minor alterations:

*Countries* selected were Bulgaria, Malta and Romania out of the EU-12. Malta is characteristically different from the other two countries, while Bulgaria and Romania has many in common at first sight. However, as a more in-depth look disclosed, many features make them distinct, despite they joined the EU at the same time in 2007. These three countries' mid-term programme evaluations will be used for international comparison of pre-selected methods (and testing of H3-5).

From the *programme* aspect, for assessing the preselected evaluation methods, one NSRF (with regional focus), two economic competitiveness OPs, a regional OP and a social OP have been selected. It contains (in the appropriate ratio) sectoral and regional programmes, NSRFs and OPs and hard (competitiveness, infrastructure, regional development) and soft (social, healthcare, education, employment) OPs, which is

beneficial for testing H3-5 (methods). More information on the programmes (subjects of the evaluations) are available later in this section.

From an *evaluation* perspective, the selection contains the final versions of 2007-13 mid-term evaluations that are publicly available in English language. It promotes easy reference, comparability and decreases the risk of misinterpretations. This way data protection and potential issues of intellectual property have also been covered. The next table (Figure 17) summarises the results of the selection process in terms of countries, programmes, evaluations and methods.

**Figure 17 – Selection result overview table (by country, programme and evaluation) for H 3-5**

Country	Programme	Evaluation	Method
<b>Romania</b>	NSRF	Mid-term <sup>10</sup>	SPOT
	SOP IEC	Mid-term <sup>11</sup>	
<b>Bulgaria</b>	OPRD	Mid-term <sup>12</sup>	SPOT, LTA
<b>Malta</b>	OPI	Mid-term <sup>13</sup>	SPOT, PCA
	OPII	Mid-term <sup>14</sup>	PCA, LTA

Source: own construction

SPOT: Simple Progress Overview Tool, PCA: Post-Contractual Assessment, LTA: Lead Time Assessment

The five programme evaluations selected for further research are the following:

*Romania: SOP IEC mid-term evaluation (2009-10) and NSRF mid-term evaluation (2009-10)*

The mid-term evaluation of the SOP IEC (economic competitiveness OP) in Romania took place in 2009-10 using methods, tools and techniques collected by the international consortium (Romania, Hungary, Scotland, Ireland, Lithuania) and the pool of experts conducting the engagement. The evaluation ran parallel to the evaluation of the NSRF in Romania, and has several inter-linkages to that evaluation. In terms of methodology used, the evaluations are fairly similar, therefore, SOP IEC and NSRF mid-term evaluations provide a sound basis for comparison of the results of identical methods in different context and scope.

<sup>10</sup> ACIS, 2010

<sup>11</sup> ACIS, 2010

<sup>12</sup> MRDPW, 2011

<sup>13</sup> PPCD, 2010a

<sup>14</sup> PPCD, 2010b

*Bulgaria: OPRD mid-term evaluation (2010)*

The mid-term evaluation of OPRD (regional and territorial development OP) in Bulgaria was carried out in 2010 by an international joint venture. The evaluation used methods and techniques similar to the Romanian mid-term evaluations. One of the reasons behind using a similar approach and toolset is that Romania and Bulgaria were the two new joiners of the EU in 2007, with no direct previous experience in using Structural and Cohesion Funds support.

*Malta: OPI and OPII mid-term evaluation (2010-11)*

Malta is a very small country, an island, whose intervention set is considerably different from CEE countries. However, the mid-term evaluations conducted in 2010-11 used techniques previously used in the CEE which provides excellent grounds for comparisons. The mid-term evaluation of OPI contains the assessment of the ERDF and CF co-funded interventions in Malta while OPII includes the ESF interventions.

The result of the selection exercise from a method point of view is the following:

**Figure 18 – Selection result overview table by evaluation method (testing H3-5)**

Method	Country	Programme	Evaluation
<b>SPOT</b>	Romania	NSRF	Mid-term
		SOP IEC	Mid-term
	Malta	OPI	Mid-term
	Bulgaria	OPRD	Mid-term
<b>PCA</b>	Malta	OPI	Mid-term
	Malta	OPII	Mid-term
<b>LTA</b>	Bulgaria	OPRD	Mid-term
	Malta	OPI	Mid-term

*Source: own construction*

The description of the methodologies and their practical use will be followed by at least two real-life examples of their applications from the selected countries of Southern, Central and Eastern Europe. The sub-section for each methodology analysis parts (Sections 3.4, 3.5 and 3.6) will be closed by the findings of the analysis by methods by hypotheses. The regional programme analysis sections (Section 3.2 and 3.3) will be followed by the three methodology analysis sections (3.4, 3.5 and 3.6), using the above selected research subjects.

## **3.2 Regional programme analysis: Cohesion Policy regional results in EU-12**

I formulated five hypotheses for my research, out of which the first two hypotheses concerns the Cohesion Policy impacts on the development and regional disparities in the EU-12 and the diversity of the EU-12 Regional Operational Programs. In this sub-section I present the analysis for testing these hypotheses along with the findings of the analysis.

### **3.2.1 Cohesion Policy context in the EU-12**

In order to analyse EU-12 programmes, it is necessary to overview the country context and the fund allocation in EU co-financed programmes in the 2007-13 budgetary period. This overview will assist in positioning the various countries of the EU-12 region so that the programmes, the corresponding mid-term evaluations and the methodologies applied within can be placed in context.

The contents of sub-section 3.2.2 rely on both the KPMG report “EU Funds in CEE” which I managed, written and edited personally (Várnai, Matolcsy, Mezősi, 2009), complemented by the data and cross-checking opportunity provided by the later released, similar Commission paper “Cohesion Policy 2007-13 National Strategic Reference Frameworks” (EC, 2007) and the country level NSRF factsheets of the European Commission. All statistical figures originate from Eurostat and use 2008 figures so that reference figures correspond to the approval of the country fund allocation plans, and the initiation of the 2007-13 programme implementation.

#### **EU-12 programme overview in the 2007-13 period**

The scope of this analysis encompasses the so-called EU-12 countries (following EC definition): Bulgaria, Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia, and Slovenia. These 12 countries joined the European Union in two waves: on 1st May, 2004 the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia and Slovenia; then on the 1st January, 2007 Bulgaria and Romania acceded the EU.

The EU-12 had a total population of cca. 110 million, with GDP per capita (EUR 11,500) and outstanding average GDP growth rate (6.9%) showing a steadily increase in the new

millennium (Eurostat, 2008)<sup>15</sup>. However, there were significant differences among the countries regarding GDP per capita and GDP growth rate as two basic development indicators. The minimum GDP per capita figure belonged to Bulgaria with EUR 3,700; while the maximum figure, EUR 17,300 was produced by Slovenia. According to GDP growth rate, Slovakia bore the top spot with 10.4%, and Hungary the last at 1.3% (Eurostat, 2008).

**Figure 19 – Overview of the EU-12 countries, basic information on GDP and EU funds**

EU-12	min	max	mean	total
<b>Country information</b>				
<b>Population (million inhabitants)</b>	1 (CY)	38.5 (PL)	9.0	108.2
<b>GDP per capita (PPP, thEUR)</b>	3.7 (BG)	17.3 (SL)	11.5	na
<b>GDP growth rate (%)</b>	1.3 (HU)	10.4 (SK)	6.9	na
<b>EU Fund information</b>				
<b>EU Funds available (bnEUR)</b>	0,8 (MT)	85.3 (PL)	18.5	222.2
<b>EU Funds per capita (thEUR)</b>	1.3 (RO)	3.2 (EE)	2.4	na
<b>EU Funds / annual GDP (%)</b>	2.1 (SL)	3.4 (EE)	2.9	na

*Source: own construction, based on country NSRFs and Eurostat (2008), EU funds = Cohesion Policy funds*

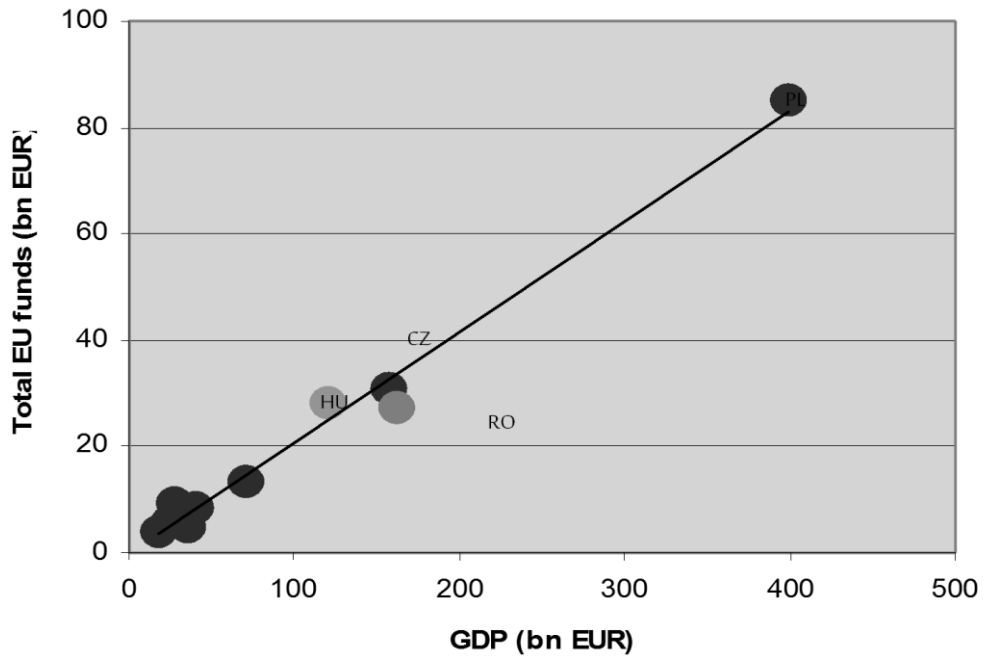
### **Breakdown of EU funds among available funds in EU-12 in 2007-13**

The total Cohesion Policy (hereinafter: EU) funds at the disposal of EU-12 countries in the 2007-13 period equalled EUR 222.2 billion, the majority (81%) of which comprised Structural and Cohesion Funds resources, while agriculture, rural development and fisheries related funds constituted around 15% of the total, the rest being split among less significant EU funding sources.

The EU-12 countries received EU co-funded budgets with respect to their GDP and population, though the final budget was always influenced by other relevant factors and the final bilateral negotiations with the EC. Looking at the results of the funds vs. GDP and funds vs. population analysis, new joiners in the 2007-13 period (Bulgaria and Romania) scored a bit below EU-12 average, while the Czech Republic, Poland and Hungary were over the average.

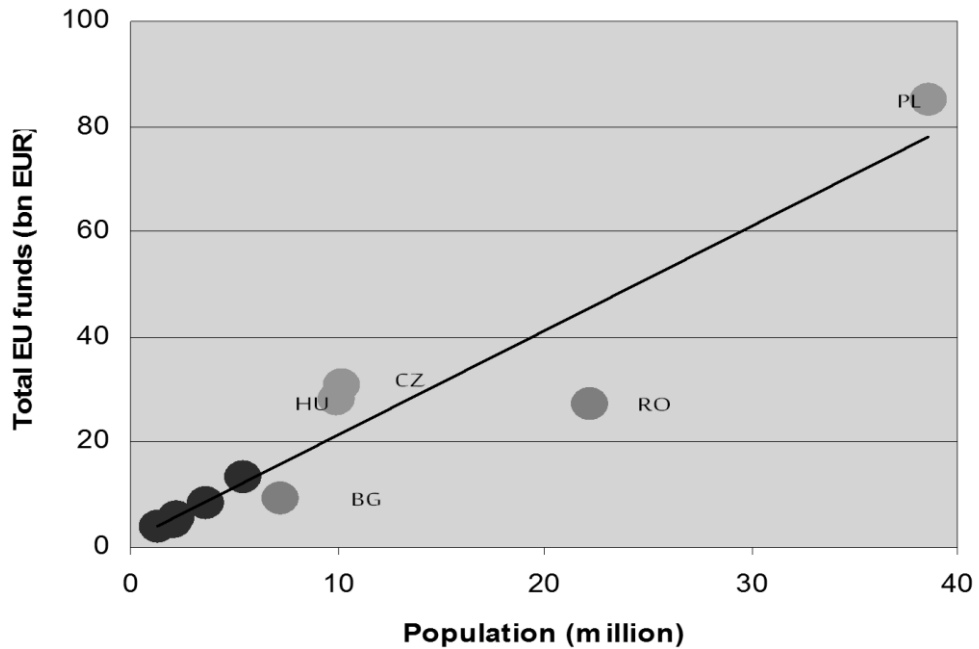
<sup>15</sup> The reason for using 2008 Eurostat data is that EU funds allocation were finalised in 2008, therefore the corresponding reference values (general social and development indicators) need to be deriving from the same point of time.

Figure 20 – 2007-13 EU-12 funds vs. GDP by country (2008 figures)



Source: own construction, based on Eurostat data (2008)

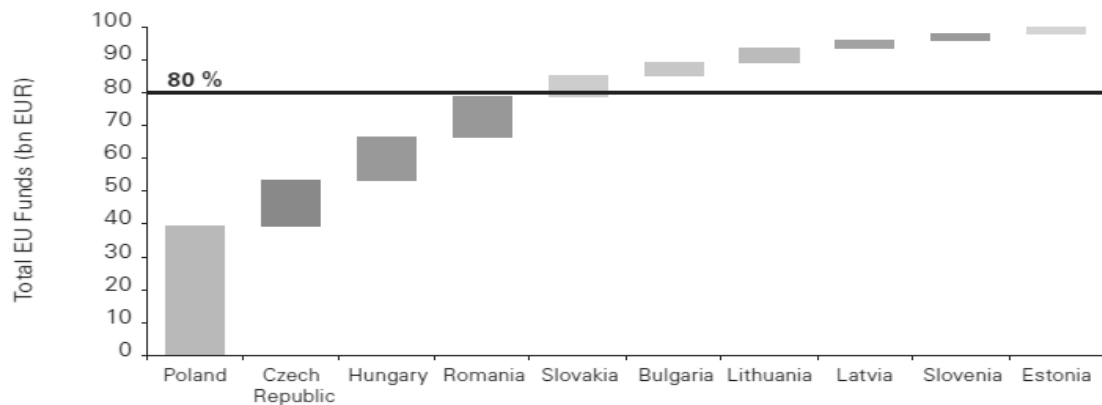
Figure 21 – 2007-13 EU funds vs. population by country (2008 figures)



Source: own construction, based on Eurostat data (2008)

It is also apparent, that the countries with the most population among the EU-12 (Poland, the Czech Republic, Hungary and Romania) accounted for almost 80% of the funds available for the EU-12 countries in this context, with the two islands, Malta and Cyprus closing the ranks.

**Figure 22 – 2007-13 EU funds breakdown by country in the CEE**



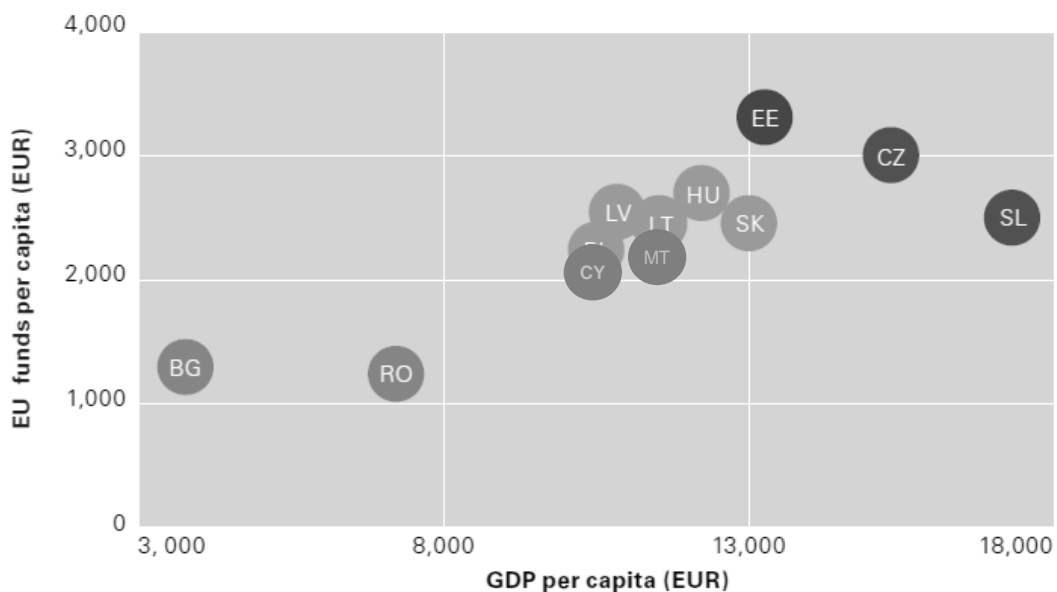
Source: own construction, based on Eurostat data (2008)

### 3.2.2 Cohesion Policy funds allocation country clustering in the EU-12

A simplified cluster analysis has been carried out in order to disclose potential relation between EU funds volume per capita versus the relative development figures (indicated by GDP per capita) of a country.

#### Cluster analysis results (based on EU funds per capita)

**Figure 23 – Clusters of EU countries – GDP per capita vs. EU funds in 2007-13 per capita**



Source: own construction, based on Eurostat data (2008)

By eliminating the influence of population from the clustering exercise, by comparing the countries based on GDP *per capita* versus EU funds *per capita*, the clustering is normalised. This result reflects the performance of the country and the corresponding EU funding volume.

In the context of the analysis, the following clusters could have been defined:

*Cluster 1 – Winner:* Estonia had the most funds at its disposal proportionally (more than EUR 3,000 per inhabitant). This enabled the country to realise more visible, tangible impacts using EU funds.

*Cluster 2 – Developing countries:* Romania and Bulgaria were catching up fast to member states joining in 2004, but their indices were still lower. The relatively low indicators of EU funds per capita and GDP per capita were balanced by the fact that these countries had been developing rapidly, with growth rates much higher than other EU members.

*Cluster 3 – Wave 2 joiners:* These Central European countries, the Baltic countries (except for Estonia, outperforming the other two) and the islands (Cyprus and Malta) were performing steadily.

*Cluster 4 – Wave 2 top performers:* These countries, namely the Czech Republic and Slovenia had been showing greater development than other Wave 2 countries.

### 3.2.3 **Regional impacts of Cohesion Policy in the EU-12 (2004-2011)**

The objective of EU Cohesion Policy is the moderation of regional inequalities. There are many authors actively researching and challenging whether Cohesion Policy proceeds along the above-mentioned objective (Riedel, 2005; Horváth 2015; Lóránd, 2009; Faragó, 2004, Ferry, 2013), also including the services of the European Union (EC, 2014d). The Cohesion Reports always provide a self-assessment on the performance of Cohesion Policy. The last, Sixth Cohesion Report states that the Cohesion Policy contributed to the GDP growth of 0.5-1.0% (EC, 2014d, p.232) in the 2007-13 period, providing evidences on its contribution to member state development. However, the report does not formulate straightforward statements on the changes of regional inequalities, though it is a fundamental role of the Policy. As I have the EU-12 countries in the scope of my research I also tested if the Policy worked adequately in this context.

There are many indicators of socio-economic development, out of which GDP per capita is the most widely acknowledged. For the quantification of gaps between regional development, there are several models and calculations used. Out of the many models, I would highlight and base my assessment on two. First, the *Williamson hypothesis* (Williamson, 1965; Nemes Nagy, 2005), describing a relation between development and homogeneity that resembles a 'U' upside down. This hypothesis claims that in the course of development of a region, first heterogeneity increases and upon reaching an inflection point, development leads to the reduction of internal development gaps. The other basis, the neoclassical theory based *beta convergence* (Major, 2001, Canova, 2004; Dvoroková, 2014) is the result of the *Solow model* (Solow, 1965) on the diminishing marginal returns on capital, claiming that underdeveloped regions grow at a rate higher than developed regions, thus catching up more quickly.

As a result of these two models, I assume (Hypothesis 1) that Cohesion Policy, aiming at regional convergence did not fulfil its role adequately: though it resulted in a faster development of underdeveloped regions, interregional inequalities in the EU-12 have actually grown. The next table shows the differences in the EU-12 group in terms of a general development indicator, GDP per capita (PPS, 100 = EU-28 average) covering the two programming periods of 2004-06 (in case of EU-10) and 2007-13 (all EU-12) together, i.e. in the 2004-2014 period.

**Figure 24 – EU-12 GDP/capita figures 2004-2014**

Countries	GDP / capita (EU-28 = 100)											Change (%)
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	
<b>Bulgaria</b>	34	36	37	40	43	44	43	44	45	45	45	<b>+32%</b>
<b>Czech Rep</b>	79	80	81	84	82	83	81	83	82	82	84	<b>+6%</b>
<b>Cyprus</b>	97	99	99	100	105	105	102	96	94	89	85	<b>-12%</b>
<b>Estonia</b>	55	60	64	69	68	62	63	68	71	73	73	<b>+33%</b>
<b>Hungary</b>	62	62	62	61	63	64	65	65	65	66	68	<b>+10%</b>
<b>Latvia</b>	48	51	55	60	60	53	53	57	60	64	64	<b>+33%</b>
<b>Lithuania</b>	50	53	56	61	63	57	60	65	69	73	74	<b>+48%</b>
<b>Malta</b>	81	81	79	78	81	84	86	84	85	86	85	<b>+5%</b>
<b>Poland</b>	49	50	50	53	55	59	62	64	66	67	68	<b>+39%</b>
<b>Romania</b>	34	35	38	42	48	49	50	51	53	54	54	<b>+59%</b>
<b>Slovakia</b>	57	60	63	67	71	71	73	73	74	75	76	<b>+33%</b>
<b>Slovenia</b>	86	86	86	87	89	85	83	83	82	82	83	<b>-3%</b>

Source: Eurostat, own construction

The diversity is obvious in trends and results across the EU-12. Though the table above does not allow drawing conclusions on the overall impacts of regional development in the EU-12 countries, it allows the formulations of opinions<sup>16</sup>. Following the phenomenon of ‘beta convergence’, Romania and Bulgaria were the most underdeveloped countries out of the EU-12 at their accession in 2007 with 40% and 42% of the EU average then. Romania indeed showed the greatest step forward with +59% and Bulgaria is also up +32%. Cyprus, the Czech Republic, Slovenia and Malta led the development ranks in 2004 and they indeed fall into the lower (if not relatively negative) development category. The beta convergence effect is underpinned by the relative development figures of the EU-12. The winners of the development race in descending order are Romania, Lithuania, Poland, Latvia, Estonia, Bulgaria and Slovakia with a development pace of more than +30% over the past decade. Hungary, the Czech Republic and Malta showed moderate increase of +5-10%, while Cyprus and Slovenia suffered a loss of -12% and -3%, respectively. On the other hand, the minimum and maximum relative GDP/capita figures of the NUTS2 level show inequalities (see next Figure).

**Figure 25 – EU-12 GDP/capita figures 2004-2011 at NUTS2 level (EU-28 = 100)**

	2004				2011			
	Country average	Regional min	Regional max	Range of regional	Country average	Regional min	Regional max	Range of regional
<b>Bulgaria</b>	34	27	51	<b>24</b>	44	29	78	<b>49</b>
<b>Czech Rep.</b>	79	62	163	<b>101</b>	83	66	171	<b>105</b>
<b>Cyprus</b>	97	-	-	-	96	-	-	-
<b>Estonia</b>	55	-	-	-	68	-	-	-
<b>Hungary</b>	62	41	101	<b>60</b>	65	40	110	<b>70</b>
<b>Latvia</b>	48	-	-	-	57	-	-	-
<b>Lithuania</b>	50	-	-	-	65	-	-	-
<b>Malta</b>	81	-	-	-	84	-	-	-
<b>Poland</b>	49	35	78	<b>43</b>	64	44	107	<b>63</b>
<b>Romania</b>	34	28	72	<b>44</b>	51	37	122	<b>85</b>
<b>Slovakia</b>	57	42	129	<b>87</b>	73	51	186	<b>135</b>
<b>Slovenia</b>	86	72	104	<b>32</b>	83	71	100	<b>29</b>

Source: Eurostat, own construction

<sup>16</sup> Please note, that as a rule of thumb, EU funding programme impacts can be measured at a minimum of 3-4 years past programme close, which poses limitations to the evaluation. On the other hand, a thorough comparison would require the setting of a „business as usual” (BAU) scenario, depicting of what could have happened without the implementation of the programme. However, as this period of time also includes the economic crisis starting in 2008, overwriting previous forecast and cycle models, such BAU scenario is not used, and I only present the changes in absolute value for reasons of simple interpretation.

The changes are dramatic: there is an overall improvement effect assumedly partly deriving from the EU co-financed interventions - similarly to other previous study findings (De la Fuente, 2003) - however, the development inequalities within countries have grown enormously (some relative gaps even doubled) between 2004 and 2011. This phenomenon also underpins the Williamson hypotheses, as in relatively more developed countries (the Czech Republic, Hungary, Slovenia), the development gap between regions did not increase at the pace of the huge leap of figures in the less developed countries (Bulgaria, Romania, Poland, Slovakia). The next table summarizes the changes in country level development figures and regional gap changes in the 2004-2011 period in the countries of the EU-12 with more than one NUTS2 level territorial unit.

**Figure 26 – EU-12 NUTS2 level development gap figures 2004-2011 (EU-28 = 100)**

	2004		2011		Change: country average (2004-11)	Change: regional range (2004-11)
	Country average	Range of regional (max-min)	Country average	Range of regional (max-min)		
<b>Bulgaria</b>	34	24	44	49	+10	+25
<b>Czech Rep.</b>	79	101	83	105	+4	+4
<b>Hungary</b>	62	60	65	70	+3	+10
<b>Poland</b>	49	43	64	63	+15	+20
<b>Romania</b>	34	44	51	85	+17	+41
<b>Slovakia</b>	57	87	73	135	+16	+48
<b>Slovenia</b>	86	32	83	29	-3	-3

*Source: Eurostat, own construction*

As a conclusion, both the Williamson hypothesis, and the attributability of the beta convergence phenomenon could have been identified in the topic of development of EU-12 countries and regions in the 2004-2011 period. However it also indicates that though Cohesion Policy has indeed supported development in these countries, but it also seem to have failed in mitigating the increasing regional polarisation within the EU-12 countries. There are multiple driving factors of the above identified and articulated differences of the growing regional disparities for instance the level of centralisation / decentralisation of the administration of a member state (Pálné, 2009), the behaviour of market players, FDI, local regulations, which all play determinant role in the overall development perspectives of a country and its programming approach, stemming partly from past programme heritage.

### **3.3 Regional programme analysis: ROPs and ROP evaluations in the EU-12**

#### **3.3.1 Regional development in the EU-12**

Countries of the EU-12 show great diversity, similar to the composition of the EU-28 (Horváth, 2015). The reason behind the diversity lies in the fact, that in the late 19<sup>th</sup> century, most of these countries (Cyprus and Malta being exceptions) belonged to different empires, i.e. the Russian Empire, the German Empire and the Habsburg Empire. After the Triannon Treaty, the composition of countries changed fundamentally: countries lost territories or gained new ones. In either case, the previously existing vertical integrations, networks have been abolished and new ones have to be created. Focus was on unification and integration and not development. With the coming of the socialist era after World War II, a slow social development started with central intentions to equalise. Regional disparities indeed decreased, however, forced industrialisation and urbanisation did not follow an evolutionary path, therefore gains remained low. Even in the 1990s, with the coming of the change of systems in these countries, local governments were in focus, while the regional levels could not build up.

As a result, at the time of EU accession (i.e. 2004 and 2007), the EU-12 countries were rather underdeveloped, usually polarised with one more developed capital and underdeveloped countryside, with a regional institutional system that formally fulfilled the EU requirements on NUTS2 regions. Though regional development acts and regulations were in force since 1996-99 in these countries (except for Poland and Slovakia) their achievements were questionable and heterogeneous (Horváth-Lóránd, 2012; Pálné, 2009).

The real change in regional development could start after the EU accession as dedicated funds were at disposal to spend and utilise at the regional level. There were many obstacles from an immature regional institutional system at the beginning of the learning curve on how to plan, develop and implement EU funded programmes and projects and there were also central planning interests taking part in this zero-sum situation in pursuing funding opportunities. In each country, sectoral operational programmes and regional operational programmes (ROPs) or regional development measures have been designed. The approach and set-up of these interventions were fairly different. ROPs and ROP

related measures (in countries without dedicated ROP) accounted for 38.5MEUR, which is around 15% of the total NSRF in EU-12.

It is a rather interesting topic to discuss, if such a low ratio of funds allocated to regional development is capable of fulfilling the mission of Cohesion Policy, i.e. decreasing regional disparities across regions in Europe. The fact that such a ratio of funds have been programmed by EU-12 member states and have also been approved by the European Commission suggests, that the absorption capacities (the ability to use EU funds for the pre-set objectives) in the regions have been seen as a risk of later funds decommitment.

The Figure below shows that countries with the highest “ROP budget / NSRF budget” ratio are the ones with the highest level of development in the group (Slovenia – 34%) and with well-funded and operated regionalised character (Poland – 23%, which after the 1998 changes introduced a voivodship system, with 16 voivodes, which fill the regional level with planning and coordination authority, with funds deriving from Structural Instrument resources). The following table contains the summary and comparison of main relevant EU-12 European member state regional (operational) programme features in the 2007-13 programming period.

**Figure 27 – EU-12 country regional operational programme and intervention details**

<b>Countries</b>	<b>No. of ROPs (#)</b>	<b>Total ROP (measure) budget (M EUR)</b>	<b>ROP (measure) / NSRF budget ratio (%)</b>
<b>Bulgaria</b>	1	1,361	<b>15%</b>
<b>Czech Republic</b>	9	4,657	<b>15%</b>
<b>Cyprus</b>	* 0	54	<b>9%</b>
<b>Estonia</b>	** 0	355	<b>9%</b>
<b>Hungary</b>	7	4,305	<b>15%</b>
<b>Latvia</b>	*** 0	263	<b>5%</b>
<b>Lithuania</b>	**** 0	755	<b>9%</b>
<b>Malta</b>	***** 0	165	<b>19%</b>
<b>Poland</b>	16	19,717	<b>23%</b>
<b>Romania</b>	1	3,726	<b>14%</b>
<b>Slovakia</b>	1	1,445	<b>11%</b>
<b>Slovenia</b>	1	1,710	<b>34%</b>
<b>TOTAL</b>	<b>46</b>	<b>38,513</b>	<b>-</b>
<b>MEAN</b>	<b>3,8</b>	<b>3,209</b>	<b>15%</b>

Source: own construction

\* Part of Development of Sustainable Communities Priority and Regeneration of Urban and Rural Areas

\*\* Part of Development of the Living Environment OP

\*\*\* A priority of Infrastructure and Services OP

\*\*\*\* A priority of Promotion of Cohesion OP

\*\*\*\*\* A priority of OPI

There have been 3 models used in the 2007-13 period to approach regional development aspect:

- 1) *One regional operational programme per region* (Hungary, Czech Republic, Romania, and Poland): This is the feature of larger and regionally more developed countries in the EU-12, where there are inherent, traditionally major differences between regional characteristics that need to be handled through Cohesion Policy instruments.
- 2) *One regional operational programme for the entire country* (Bulgaria, Romania, Slovakia, and Slovenia): Those chose this option where there is a clear intention not to diversify development but to centralise distribution of funding. In case of Slovenia, it is rather the country size that lead to the elaboration of a single ROP.
- 3) *Regional development part of an other operational programme* (Estonia, Latvia, Lithuania, Malta and Cyprus): the Baltic states and Malta are too small in territory and population (they all have SPDs instead of NSRFs) to need and to have separate, dedicated regional operational programmes.

**Approach to funds allocated to the regional aspect of development is fairly different among the EU-12 countries.**

- 1) *Lower spending on regional development* (5-9-11% of total EU funds): Baltic States, Cyprus and Slovakia.
- 2) *Moderate allocation on regional development* (14-15% of total EU funds): Bulgaria, the Czech Republic, Hungary and Romania
- 3) *Highest allocation on regional development* (19-23-34%): Malta, Poland, Slovenia.

**The overall concentration of direct regional development funds is high among the EU-12 countries.** Poland in itself accounts for 51% of regional development funds allocation in the EU-12, while the next three countries (the Czech Republic, Hungary and Romania, in this order, with 12%, 11% and 10%) cover another third of it. The share of the remaining 8 countries altogether is less than 16%.

### 3.3.2 **ROP measure structure and content level comparison in the EU-12 countries**

The structure and content is similar in the 2007-13 EU-12 regional programmes.

ROP measures typically intend to co-finance improved accessibility (including local transport – roads, public transport, bicycle lanes, etc.), improved tourism potential, local infrastructure (social, educational, public spaces) and environment (rehabilitation areas, contaminated areas). In some cases (Romania, Slovenia, Hungary) elements of local economic development appear.

The intervention objectives of the ROPs and measures build on the development gap issues that are the results of socio-economic heritage originating from late 19th century differences, and uniformisation of the soviet era. The interventions usually target operational level objectives (means), but only indirectly pursue strategic level goals like extended regionalism, more authority to regions or strengthening the second level of cities, i.e. (potential) regional centres to ease the polarised situation in these countries governed by the overwhelming power of the centre, with unmatched development differences compared to the countryside.

The balance and relative weight of these measures largely depend on the current assets, strengths and weaknesses, needs of the respective regions and the future vision for development.

The next table provides a one-page summary of the intervention objectives and contents of the EU-12 regional programmes in 2007-13.

**Figure 28 – Intervention structure and content of EU-12 regional development measures**

<b>Countries</b>	<b>Priorities and focus areas</b>	<b>Countries</b>	<b>Priorities and focus areas</b>
<b>Bulgaria</b>	<ul style="list-style-type: none"> <li>• Sustainable and integrated urban development</li> <li>• Regional and local accessibility</li> <li>• Sustainable tourism development</li> <li>• Local development and co-operation</li> </ul>	<b>Romania</b>	<ul style="list-style-type: none"> <li>• Development of local and county road network, regional airports and ports;</li> <li>• Development of and equipping of the health services;</li> <li>• Development of equipping of social services infrastructure;</li> <li>• Development of equipment of public safety interventions in emergency situations</li> <li>• Development of pre–university and continuous vocational training infrastructure;</li> <li>• Development of business support structures;</li> <li>• Industrial sites rehabilitation;</li> <li>• Support to develop micro-enterprises;</li> <li>• Rehabilitation and sustainable valorisation of cultural &amp; historical heritage and setting up &amp; modernization of related infrastructure;</li> <li>• Creation / development / modernization of the specific infrastructure for sustainable valorisation of natural resources with tourism potential;</li> <li>• Development of accommodation and related utilities, leisure tourist infrastructure;</li> <li>• Integrated urban development plans;</li> <li>• Support for the implementation, management and evaluation of the ROP;</li> <li>• Support for the publicity and information activities of the ROP.</li> </ul>
<b>Czech Republic</b>	<ul style="list-style-type: none"> <li>• Urban regeneration and development</li> <li>• Integrated support of local development</li> <li>• Transport accessibility and services</li> <li>• Sustainable development of tourism</li> </ul>		
<b>Cyprus</b>	<ul style="list-style-type: none"> <li>• Integrated urban development</li> <li>• Regional accessibility</li> </ul>		
<b>Estonia</b>	<ul style="list-style-type: none"> <li>• Integral and balanced development of regions</li> </ul>		
<b>Hungary</b>	<ul style="list-style-type: none"> <li>• Regional economic development</li> <li>• Tourism-related developments</li> <li>• Transport infrastructure</li> <li>• Human infrastructure</li> <li>• Settlement development</li> </ul>		
<b>Latvia</b>	<ul style="list-style-type: none"> <li>• Polycentric development</li> </ul>		
<b>Lithuania</b>	<ul style="list-style-type: none"> <li>• Local development</li> </ul>		
<b>Malta</b>	<ul style="list-style-type: none"> <li>• Urban regeneration and improving the quality of life</li> </ul>	<b>Slovakia</b>	<ul style="list-style-type: none"> <li>• Infrastructure of education</li> <li>• Infrastructure of soc. Services, soc.-legal protection and soc. guardianship</li> <li>• Strengthening cultural potential of regions &amp; development of tourism</li> <li>• Regeneration of settlements</li> <li>• Regional communications ensuring transport serviceability of regions</li> </ul>
<b>Poland</b>	<ul style="list-style-type: none"> <li>• Growth of regional enterprise competitiveness</li> <li>• Increase of regional information society</li> <li>• Local transport infrastructure development</li> <li>• Environmental and ecological safety</li> <li>• Environmental friendly energy infrastructure</li> <li>• Utilization and promotion of tourism and culture</li> <li>• Modernization and development of education</li> <li>• Modernization of health infrastructure</li> <li>• Renovation of degraded city area</li> </ul>		

Source: own construction, based on country Regional Operational Programmes and direct regional development measures

### 3.3.3 ROP mid-term evaluation analysis in the EU-12

The mid-term evaluation assessment of the EU-12 ROP measures provides an external view and input to the research.

For this analysis I chose relevant regional operational programmes and the programmes containing regional development interventions of Romania, Bulgaria and Malta the pre-selected countries (see Section 3.1). Out of the five mid-term evaluations that I also used for the methodology analysis (assessment of evaluation methodologies, see Sections 3.4, 3.5 and 3.6) I chose three, from three different countries to assess their ROP's general alignment to Community expectations in terms of focal points to their respective mid-term evaluations:

#### *Romania: NSRF mid-term evaluation (2009-10) / ROP evaluation*

The mid-term evaluation of the NSRF in Romania took place in 2009-10 using methods, tools and techniques collected by the international consortium (Romania, Hungary, Scotland, Ireland, and Lithuania) and the pool of experts conducting the engagement. The only programme evaluation that was conducted prior to NSRF evaluation was the mid-term evaluation of the Regional Operational Programme in Romania, therefore, the NSRF evaluation could make use of its findings and conclusions.

#### *Bulgaria: OPRD mid-term evaluation (2010)*

The mid-term evaluation of OPRD (regional and territorial development OP) in Bulgaria was carried out in 2010 by an international (Bulgarian, Russian and Hungarian) joint expert group. The evaluation used methods and techniques similar to the Romanian mid-term evaluations. One of the reasons behind using a similar approach and toolset is that Romania and Bulgaria were the two new joiners of the EU in 2007, with no direct previous experience in using Cohesion Policy support other than the similar IPA.

#### *Malta: OPI mid-term evaluation (2010-11)*

The mid-term evaluations conducted in 2010-11, used tool and techniques previously used in the CEE which provides excellent grounds for comparisons. The mid-term evaluation of OPI contains the assessment of the ERDF and CF co-funded interventions in Malta, accounting for regional development related interventions as well.

## Analysis set up

It is one of my research hypotheses (H2) to assess ROPs in the EU-12. As many of these programmes already had a mid-term evaluation ready and available, I also used them as inputs. I started by assessing them against the general EC expectations regarding mid-term evaluation focus areas and their intensity as set out in the respective methodological working document (EC, 2006a).

The next table shows a summary on the focus of the ex ante, mid-term and ex post evaluations in the 2007-13 period (based on the Working Document No. 1: Indicative Guidelines on Evaluation Methods: Ex-Ante (EC, 2006a), complemented by own relevant aspects). This table will be used in the assessment as the reference point for comparison.

**Figure 29 – Focuses of evaluation stages in the 2007-13 programming period**

Stage	Ex ante	Mid-term / on-going	Ex post
<b>Evaluation focus</b>	Planning	Implementation	Accountability
<b>Timing</b>	2006	2007-15	2015+
<b>Rationale</b>	***	*	*
<b>Relevance</b>	***	***	*
<b>Effectiveness</b>	*	***	**
<b>Efficiency</b>	*	**	**
<b>Impact</b>	na	*	***

Source: EC, 2006a

Legend: \* - not in focus, \*\* - normal focus, \*\*\* - strong focus

The stars in the table above stand for the strength of the focus in the evaluation stages. This shows the clear intention of the Commission to have the member states concentrate on aspects and topics relevant to EC interests. However, in the 2007-13 period it was the task and responsibility of member states to commission and conduct programme evaluations, while methodology and thematic guidance was put in place by the Commission, however, not being compulsory.

## Findings of the analysis

I have compared each aspect of the above Figure (EC, 2006a) with the contents of the selected three mid-term evaluations, to find the following.

### *Findings by focus areas*

The evaluation of *rationale* was missing in all three evaluation cases. However, it is important to note that rationale and relevance are closely linked, and follow each other generally. Programme rationale should not normally be a question in mid-term, however, challenging relevance also entails questioning the rationale behind the programme. According to my analysis, evaluations of rationale in the selected mid-term evaluations blended in evaluations of relevance. This is a rather methodological question whether to split or merge these two, however, in my view, evaluation of rationale did not appear in evaluations as a result of low EC expectation and not as a result of non-relevance. In fact, all four mid-term evaluations had elements of rationale evaluation, all covered under relevance evaluation.

*Relevance* of the programmes was stressed in all three mid-term evaluations. The focus of the evaluation of relevance was continued relevance of the programme, questioning whether broader country context in the time of planning, on which the programmes have been built, have changed to the extent that requires modifications in programme content or delivery mechanisms. In all three programmes, relevance was a definite focus, and EC had also put relevance priority to “strong focus”. The EC expectation and the outcome met in this case as it was a joint interest of the EU and the member states to utilise maximum funding which required clear targeting and underlying needs analysis.

Evaluation of *effectiveness* was expected to have strong focus (according to EC, 2006a), and it actually constituted a major topic in all three evaluations. In fact, the potential for disclosing real effectiveness issues is limited in mid-term, usually examined through time-proportional progress (in terms of indicator values and financials). However, progress indicators of EU co-funded programmes cannot be characterised by linear tendencies (e.g. spending is not necessarily time-proportional). All three evaluations have put emphasis on progress evaluations, showing that EC and member states’ interest were identical in this view, i.e. a clear picture on progress status, odds of achieving pre-set objectives, and feedback on what works and what needs improvement (better work or better target).

Evaluation of *efficiency* is a normal focus according to the working document, however, in all three mid-term evaluation cases, this was top priority. The answer to this discrepancy lies in four main reasons: First, there is a clear member state interest to spend

funding for the right purpose, in an accountable way. The success of all programme implementation aspects relies on the institutional system, on their capacity and competence to plan and execute. If the institutional system is not efficient, it indeed has a hindering effect on both relevance, effectiveness and potential impact. Second, EU-12 countries have not had a long track record of implementing EU funded programmes, therefore they were in the early stages of the learning curve in terms of internal competence and efficiency and more than eager to learn what they did well and what needed improvement. Third, competence was limited in the respective country labour markets as well (number of persons having acquired experience in EU funds implementation was still low proportionally (Bulgaria, Malta) and in absolute value (Malta) leaving little room for quick quality changes in competence. Fourth, institutional structures were driven by political factors (Bulgaria, Romania) or were still under change (Romania, Malta) and feedback was awaited for current performance levels.

Evaluation of *impact* was not explicitly part of the three mid-term evaluations. However, evaluation of effectiveness had to provide an insight to what impacts can potentially be expected, based on time-proportional progress and extrapolation of results achieved. It is too early to assess impacts in mid-term which creates an objective barrier to evaluating it. In this view, it is rather questionable why EC had put “not in focus” in mid-term instead of the “na” that is represented in the ex ante phase for obvious reasons.

#### *Summary of findings*

The focus of the evaluations was more intense than the EC had outlined. Evaluation of rationale and impacts were moved to the back, while relevance, effectiveness and efficiency were all highlighted topics.

In fact, three out of the five evaluation aspects were different from that of the guidance. I have to note, however, that these changes are not dramatic as there is only one step forward or backward on the four-level scale (from 0 to 3 stars in intensity). What is more dramatic, however, is the narrowing of the scope of the research from five to three topics, as two of them (rationale and impact) have been practically ignored in the course of the evaluation.

I want to emphasize that the difference is not the fault of the evaluators, who carried out the tasks. Having checked the ToRs of the evaluations in the sample, it is clear that the

evaluators followed the path outlined by the commissioners of the evaluation, as the ToRs were previously elaborated in a way that overruled the EC guidance.

The summary of findings is presented in the table below. The first column contains the expectations (EC, 2006a), while the second summarises findings of the analysis.

**Figure 30 – Summary of analysis findings (ROPs vs. EC Working Document No.1)**

Stage	Mid-term / on-going EXPECTATION	Selected OPs IN FACT	Comment on DIFFERENCE
<b>Rationale</b>	*	na.	-
<b>Relevance</b>	***	***	0
<b>Effectiveness</b>	***	***	0
<b>Efficiency</b>	**	***	+
<b>Impact</b>	*	na.	-

Source: EC, 2006a

Legend: \* - not in focus, \*\* - normal focus, \*\*\* - strong focus  
+ more done - less done 0 - no change

EU Commission and member state programmers and evaluators should learn the lessons from the alteration of ToRs and evaluations relative to the EC methodology guidance. It is important to clarify whether relevant EU expectations are strict and should be followed by the letter (and in this case, more consultations are required at member state level to find a viable compromise) or the EC maintains the role of guidance with great freedom to the member states on interpretation (however, it might lead to inadequate answers to EC evaluation questions).

The fact that the framework of programming became significantly stricter and more rigid for the 2014-20 period, shows the will of centralisation and more control on funds allocation and spending on EC side.

## **3.4 Methodology analysis: Simple Progress Overview Tool (SPOT)**

### **3.4.1 Method overview**

#### **Financial progress**

One of the most obvious starting points for mid-term evaluations is identifying the current status of progress, including financial progress. The issue of financial progress is in strong coherence with all three evaluation themes of the mid-term evaluations, i.e. relevance, efficiency and effectiveness and, therefore, it is in the core of setting the scenes for evaluation.

In a nutshell, in terms of relevance, financial progress indicates whether the programme follows the right set of goals driven by the socio-economic environment. As for efficiency, financial progress points at two areas: how the institution system copes with implementation (i.e. grant allocated / cost of implementation ratio) and the cost-effectiveness of the programme (i.e. grant allocated / programme objective indicator ratio), where both ratio uses current level of grant allocated as the numerator. Effectiveness assesses the current level of financial progress vs. programme objectives. Please note that “allocation” of grant stands for grants “reimbursed” or “spent” or “paid” in this context, and not referring to initial financial allocation or “setting the budget” for the programme.

#### **Quantification of financial progress**

Progress is one of the most relevant attributes of any programme implementation. One of the first questions of the evaluation of progress is „Where are we now?“ or „How much have we spent so far?“. First, an indicator has to be defined that can individually, attributably and precisely interpret the current financial progress level. This value – by nature of grant programmes – is absorption that we can define as grant spent per budget allocated to the programme. Therefore, progress is measured by the grant amount already „spent“ on programme beneficiaries. In order to assess if spending is on time or behind schedule, the current level of absorption has to be compared with the planned scheduling of spending. In case there is no plan, the simplest estimate to assess progress is to assume linear spending by using total programme budget / number of programme years as a reference base to estimate the expected annual spending level. Please note that evaluations also concern progress in terms other than financial progress. Progress in

general is usually measured through the system of programme/impact/result/output indicators. In the frameworks of the current study, „progress“ always mean „financial progress“.

### **Project selection process**

As a first step to understanding financial progress of a grant programme, the grant allocation process (e.g. an application process or other type of project selection process) has to be clearly defined. In order to demonstrate the use of SPOT methodology, a standard grant application process will be assumed, consisting of a three-level assessment system including an administrative, an eligibility and a technical assessment on the application submitted. In this context, it is presumed that if an application passes through all three levels, then it is approved, a grant contract is signed and payment can be initiated to the beneficiary of the grant.

A standard application process for demonstrating purposes (that is very similar to those often used in National Strategic Reference Frameworks in the 2007-13 EU programming period) includes ten steps from submission of application up to payment of the grant. Please note that in some of the member states the three-stage process is reduced to two or even one in practice by integrating one stage into another. This process is depicted in the next figure.

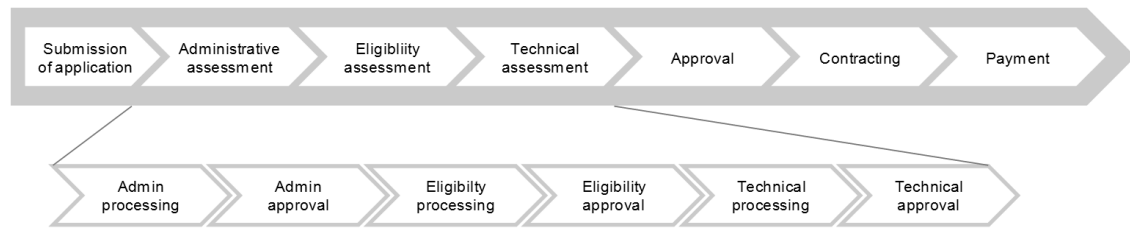
**Figure 31 – A standard project selection (application) process using a three-stage selection system**



*Source: own construction*

In order to measure financial progress, one has to consider all the above-mentioned stages and the corresponding grant amounts. Also, one has to consider that those stages of a project selection or application system that include assessment, have two factors linked to them. The first one is the “process” aspect, indicating if the application covering a certain grant amount has been processed at that stage. The second one is the “pass” aspect, clarifying if the application carrying the grant has been approved to pass to the next level of the project selection process. Therefore, all assessment phases has to be decomposed to a “processed” and a “passed” element.

**Figure 32 – Breakdown of the three-stage selection system to processing and approval elements**



*Source: own construction*

### **The SPOT tool**

The method and tool described in this dissertation is a general one, capable of over-viewing financial progress of a programme, thus supporting programme evaluation activities.

The Simple Progress Overview Tool (SPOT) is a tool used to identify the reasons underlying a figure describing, at a predefined point of time, the level of absorption (i.e. the value of reimbursed grants) of an intervention by breaking that figure down to constituent, individually meaningful factors that (are likely to) have a direct, substantial impact on the final result.

SPOT factors are individually meaningful numerical indicators generated and applied in the course of the analysis of the value which is expressed in the form of ratios illustrating significant stages in the process of project selection. Multiplication of the SPOT ratios for an intervention will equal absorption ratio.

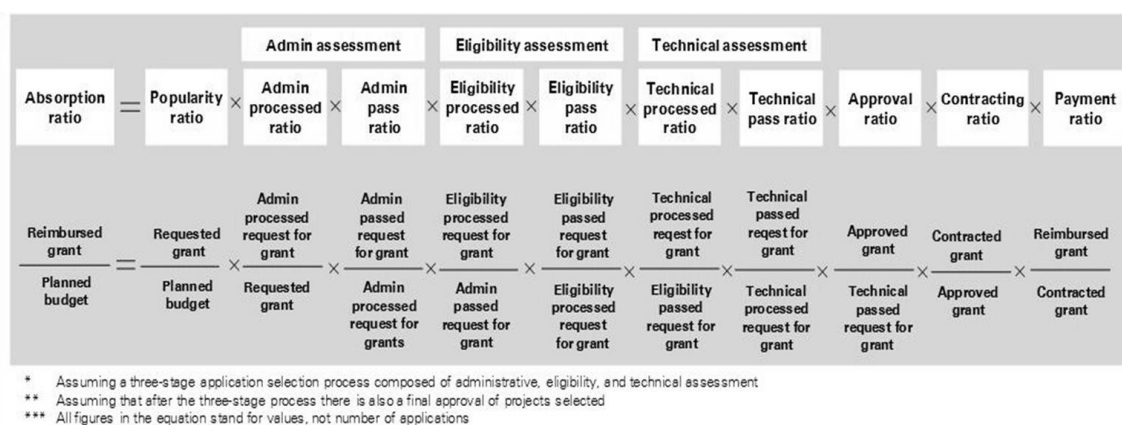
The final output of the SPOT analysis can also create a solid basis for estimating absorption perspectives of the programmes by building a simple mathematical model on the extrapolation of factors and defining potential scenarios (without intervention, or with intervention to one or more than one factors).

### **The SPOT equation**

The basis of the SPOT methodology and tool derives from the definition of absorption. In mathematical terms, the absorption ratio is the value of reimbursed grant amount divided by the planned budget of the programme. The absorption ratio clearly indicates how much grant the programme has spent to date.

The absorption ratio can be divided into several other multiplying factors (ratios) by bringing in the financial values corresponding to the pre-defined stages of the project selection process. Following the Du Pont approach, the mathematical product of the individual ratios equals the absorption ratio after simplification. The next figure summarises the “SPOT equation”, the basic concept of the tool.

**Figure 33 – SPOT equation tool (assuming a three-stage project selection system)**



Source: own construction

By the de-composition of the mathematical product, the following ratios can be obtained:

- **Popularity ratio** = *requested grant / planned budget*.  
This ratio describes the level of demand for the selected intervention.
- **Admin processed ratio** = *admin processed requested grant / requested grant*.  
This ratio describes the proportion of requests already processed by the project selection system.
- **Admin pass ratio** = *admin passed request for grant / admin processed request for grant*  
This ratio describes the proportion of requests that received positive feedback and have passed the first level of assessment.
- **Eligibility processed ratio** = *eligibility processed request for grant / admin passed request for grant*  
This ratio describes the proportion of administratively adequate requests that has already been processed by the project selection system.
- **Eligibility pass ratio** = *eligibility passed request for grant / eligibility processed request for grant*

This ratio describes the proportion of administratively adequate requests that received positive feedback and have passed the second level of assessment.

- **Technical processed ratio** = *technical processed request for grant / eligibility passed request for grant*

This ratio describes the proportion of requests that have met administration and eligibility criteria and have been processed by the project selection system.

- **Technical pass ratio** = *technical passed request for grant / technical processed request for grant*

This ratio describes the proportion of requests that requests that have met administration and eligibility criteria and received positive feedback in terms of technical assessment and therefore have passed all three levels of assessment.

- **Approval ratio** = *approved grants / technical passed request for grant*

This ratio describes the proportion of requests that have been found adequate in terms of both administrative, eligibility and technical criteria and have been approved by the respective authority. This ratio fundamentally reflects a “processed” element, though, in some rare cases might carry a “passed” element.

- **Contracting ratio** = *contracted grants / approved grants*

This ratio describes the proportion of approved requests that have already been contracted out. This ratio is a “processed” ratio by nature.

- **Payment ratio** = *reimbursed grant / contracted grants*

This ratio describes the proportion of contracted grants that have already been transferred to the beneficiaries. This ratio is a “processed” ratio by nature.

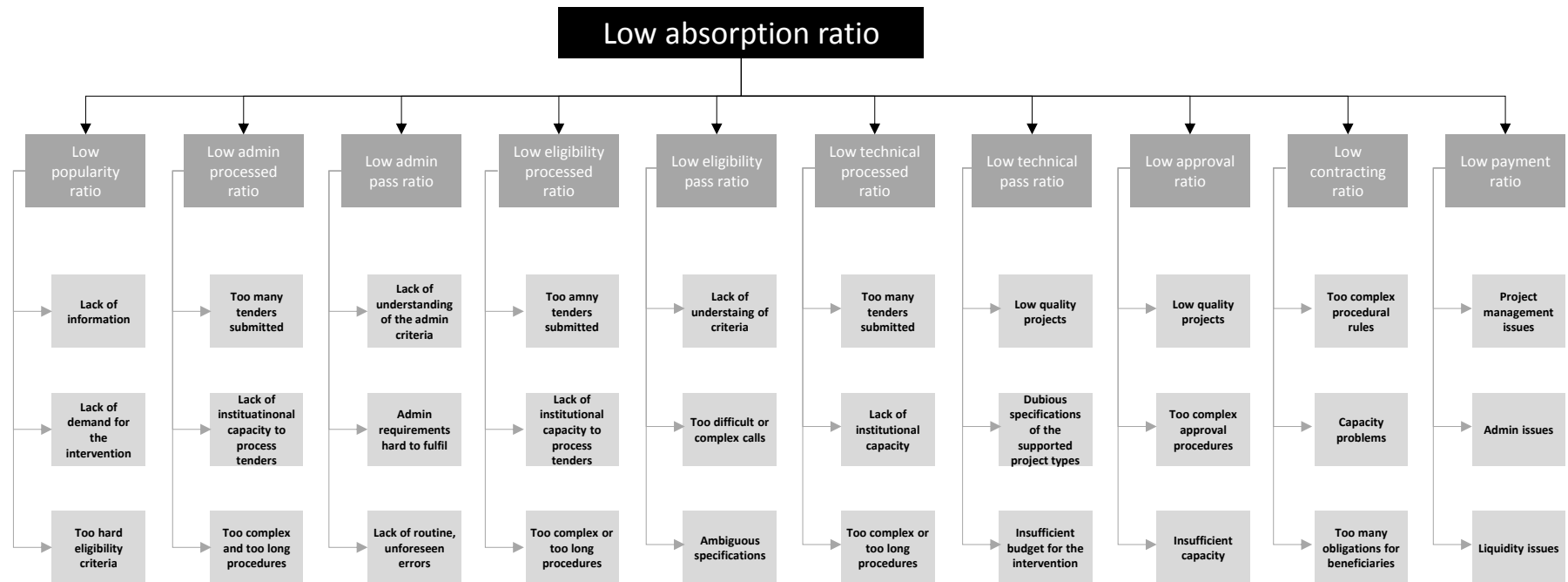
Consequently, the detailed ratios provide insight and an increased understanding of the reasons behind the overall “absorption ratio”. The level of the de-composed ratios within the absorption ratio assists the identification of potential bottlenecks in the project selection system. Examination of the more detailed ratios allows the different factors underlying relatively high or relatively low absorption ratios to be identified, and hence the nature of the problems to be explored in more detail and better targeted further assessments to be developed. In order to assess if a ratio is high or low, comparisons to 1) similar previous programmes 2) to other similar programme providers (e.g. member states) or 3) to absolute values in terms of the “processed” ratios (100% if all inputs have been processed).

As a result of the de-composition of absorption, the above-mentioned individual ratio values are calculated for the relevant programme level (i.e. in terms of EU Structural Funds and Cohesion Fund programmes they are National Strategic Reference Frameworks, Operational Programmes, measures, sub-measures). Some potential explanations (examples) to the low ratio values (constituting potential bottlenecks in the application system to be assessed in more depth) disclosed as a result of filling-in the SPOT equation with relevant data are in the next Figure (Figure 34).

It is important to point out, that low values in the equation can always be attributable to either the EU programme planning and implementation system or the beneficiary. For instance a low popularity ratio may be the result of overcomplicated or irrelevant calls for proposals, or the lack of proper communication towards the target group, or, on the other hand, lack of interest from the beneficiary side.

However, evaluators are mandated to formulate recommendations for the institutional system, therefore all bottlenecks and areas for further improvement in the system is viewed upon as an improvement gap for the institutional system. Staying with the previous example, it is the responsibility of the EU programming and implementation system to plan and publish call for proposals that are in line with the previously assessed and discussed development needs and capacities of the potential final beneficiaries. It is a major part of the planning activity to triangulate the conditions of the planned calls for proposals among limitations posed by EU regulations and policies, limitations posed by the development policy and institutional system of the Member state and development needs and capacities of the potential beneficiaries.

Figure 34 – SPOT equation tool (on a three-stage project selection system) explanations for low ratios



Source: own construction

Hypotheses developed upon the evidence based results of the SPOT equation can establish a sound basis for further investigation of reasons behind progress.

The SPOT equation and its earlier version (de-composition) has been used in many member states' mid-term evaluations in the 2007-13 period. Therefore it is a great opportunity to explore and compare how different countries used the methodology, and what results it showed.

### 3.4.2 International examples

#### International example 1 – NSRF mid-term evaluation, Romania

In the 2007-13 Romanian NSRF, there have been seven Operational Programmes (ROP, SOP ENV, SOP T, SOP IEC, SOP HRD, OP DAC OP TA). The aggregated financial figures of these programmes equal the NSRF figures.

Figure 35 – Progress factors at 30/06/2009 and 30/06/2010 by OP of the Romanian NSRF (%)

	Popularity ratio	Approval ratio	Contracting ratio	Payment ratio	Absorption ratio
<b>ROP</b>	254%	20%	90%	17%	7,9%
<b>SOP ENV</b>	82%	51%	84%	11%	3,8%
<b>SOP T</b>	66%	31%	100%	3%	0,7%
<b>SOP IEC</b>	400%	12%	70%	15%	5,1%
<b>SOP HRD</b>	189%	28%	70%	8%	3,2%
<b>OP DAC</b>	116%	25%	82%	7%	1,7%
<b>OPTA</b>	34%	76%	97%	5%	1,3%
<b>NSRF* 30/06/2010</b>	<b>168%</b>	<b>24%</b>	<b>82%</b>	<b>11%</b>	<b>3,7%</b>
<b>NSRF 30/06/2009</b>	<b>46%</b>	<b>72%</b>	<b>61%</b>	<b>18%</b>	<b>1,5%</b>

Source: ACIS, 2010

\* These figures concern the NSRF without the ETC Programmes

The analysis was not separated to “process” and “pass” factors, and the assessment stages (administrative, eligibility, technical) had to be integrated into one stage for evaluation purposes as a result of the lack of robust data and lack of consistency in project selection systems applied for each Operational Programme. As a result, the “Approval” stage covers all project selection related (i.e. administrative, eligibility, technical) ratios.

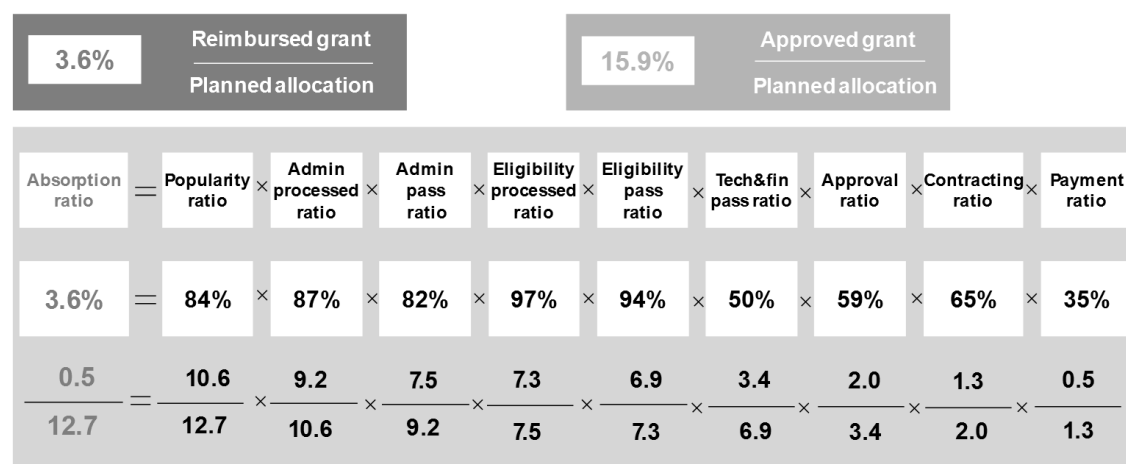
Therefore, the method could use four distinct stages to de-compose the absorption ratio into, i.e. popularity ratio, approval ratio, contracting ratio, payment ratio.

### International example 2 – SOP IEC mid-term evaluation, Romania

However, in the case of SOP IEC (which is the economic competitiveness related Operational Programme in Romania) mid-term evaluation, the “process” and “pass” factors could have been separated as a result of available data. Please note that the discrepancy in the main progress indicator (i.e. absorption ratio) compared to the SOP IEC value of the NSRF evaluation mentioned above come from the different cut-off date and data clarity issues.

The SPOT equation of SOP IEC at 30.09.2010 showed the following figures:

Figure 36 – SPOT equation of SOP IEC at 30.09.2010



Source: SOP IEC MA, 2011

The explanation briefly indicates that the outcomes of the analysis show a rather balanced picture. The OP’s popularity ratio was relatively high as a result of demand from SMEs. The application system processing and pass ratios were outstanding (very close to the maximum of 100%). Technical and financial verification together seemed to filter out half of the number of applicants. Two thirds of successful applications have signed the support contract. The payment ratio was relatively high at that time compared to other OPs and is also impressive in terms of time-proportionality (SOP IEC MA, 2010)

### International example 3 – OPI mid-term evaluation, Malta

A third example is the Maltese OPI, which is the ERDF based Operational Programme of Malta. Its range of interventions covers more than SOP IEC which was a challenge to use for SPOT.

Please note that in the case of Malta, “disbursement ratio” equals “payment ratio” of the SPOT methodology. Please also note that “process” factors are missing as a result of lack of data, and that the OP figures are broken down to the ERDF financed Priority Axis level (PA1,2,4 and 6).

**Figure 37 – SPOT equation of OPI Malta ERDF measures at 30.09.2010**

OP I - ERDF DCM Ratio Analysis					
	PA 1	PA 2	PA 4	PA 6	Fund
Popularity Ratio	154.0%	168.3%	139.76%	216.5%	172.2%
Eligibility Pass Ratio	81.9%	86.5%	81.87%	51.1%	71.7%
Approval Ratio	60.9%	41.1%	53.29%	40.7%	48.4%
Disbursement Ratio	15.3%	4.7%	0.54%	14.9%	9.1%
<b>Utilisation Ratio</b>	<b>11.8%</b>	<b>2.8%</b>	<b>0.33%</b>	<b>6.7%</b>	<b>5.5%</b>
<b>Progress Ratio</b>	<b>76.7%</b>	<b>59.8%</b>	<b>61.0%</b>	<b>45.0%</b>	<b>59.7%</b>
<b>Absorption Ratio</b>	<b>14.8%</b>	<b>4.0%</b>	<b>0.7%</b>	<b>14.5%</b>	<b>8.8%</b>

Source: own calculation, based on PPCD dataset, 30.09.2010

The next table show the Maltese OPI Cohesion Fund financed Priority Axis (PA) breakdown. Please note that in this case only the main progress indicators, i.e. progress and disbursement (absorption) ratios could have been measured due the characteristic of these PAs and the limitations of the SPOT. The limitation derives from the factor that the PAs referred below contained only a few but large infrastructure projects, which do not change incrementally over time, but drastically in a few steps. Therefore, in this case – though data was available – it was not worth to display mid-stage figures as they would have been less meaningful and comparable to other interventions, therefore less relevant to the purpose of the evaluation. These intervention require other methods and means to quantify their rate of progress.

**Figure 38 – SPOT equation of OPI Malta CF measures at 30.09.2010**

OP I - CF DCM Ratio Analysis			
	PA 3	PA 5	Fund
Progress ratio	42.05%	63.1%	52.5%
Disbursement ratio	0.05%	31.3%	18.7%
<b>Utilisation ratio</b>	<b>0.02%</b>	<b>19.8%</b>	<b>9.8%</b>
<b>Absorption Ratio</b>	<b>0.5%</b>	<b>26.9%</b>	<b>13.6%</b>

Source: own calculation, based on PPCD dataset, 30.09.2010

#### **International example 4 – OPRD mid-term evaluation, Bulgaria**

The next international example is the regional operational programme (OPRD) in Bulgaria, covering each and every Bulgarian region, consisting of regional development related interventions.

The examination of the absorption and progress ratios for the entire OPRD and for the PAs showed that the OPRD had a good popularity ratio, very high administrative and eligibility verification pass ratios, an appropriate approval ratio, a relatively high contract ratio and a moderate payment ratio. The values for each ratio are presented in the next Figure.

Generally, the ratios were adequate or good, compared to international benchmarks and achievable scores (as in the case of process ratios the figure cannot exceed but should converge to 100%). There were three outstanding values:

*Progress in Tourism (PA3):* The Tourism priority and its interventions needed further improvement regarding progress ratio (15% and 32% time-proportionally to the period 2007-10) of the initial budget allocation. The fact that the call could not start well constitutes a potential risk for absorption in this phase of project implementation (i.e. more than half of the implementation period was over at that time). However, there have been several calls launched in the meantime with the potential to promote absorption.

*Low contracting ratio in Urban development (Operation 1.1):* The figure of 46% shows that less than 5 out of 10 approved projects have been contracted. This in fact was merely an administrative issue reflecting the time requirement for signing the contract by both parties. The low rate of contracting might be explained by the relatively large number of approved projects on the reserve list as a result of lack of resources.

*Generally low payment ratios:* The payment ratio was 21% on average, showing that only one fifth of the contracted grants have already received payment (advance payment or reimbursement) up till the cut-off date. The only exceptions were Operation 4.1 and 4.2 (Local development) with time-proportionally high (50% and 38% respectively) figures; and Operation 1.1 (Urban development) with a moderate (25%) value.

As an overall assessment, the above-mentioned negative factors did not put the implementation of the whole OPRD at significant risk.

Figure 39 – SPOT factors of progress of OPRD Bulgaria at 31.12.2010



Priority Axis	Operation code	Operation name	Popularity ratio	Adm in pass ratio	Eligibility pass ratio	Tech&fin pass ratio	Approval ratio	Contracting ratio	Payment ratio	Allocation ratio	Absorption ratio
			%	%	%	%	%	%	%	%	%
PA1 Urban Development	1.1	Social Infrastructure	228%	91%	96%	90%	94%	46%	25%	169%	19%
	1.2	Housing	-	-	-	-	-	-	-	-	-
	1.3	Organisation of Economic Activities	-	-	-	-	-	-	-	-	-
	1.4	Improvement of Physical Environment and Risk Prevention	325%	61%	71%	78%	90%	96%	20%	98%	19%
	1.5	Sustainable Urban Transport System	-	-	-	-	-	-	-	-	-
PA2 Accessibility	2.1	Regional and Local Road Infrastructure	144%	99%	93%	95%	97%	76%	19%	122%	18%
	2.2	ICT Network	-	-	-	-	-	-	-	-	-
	2.3	Access to Sustainable and Efficient Energy Resources	-	-	-	-	-	-	-	-	-
PA3 Tourism	3.1	Enhancement of Tourism Attractions and Related Infrastructure	126%	91%	70%	91%	89%	11%	0%	65%	0%
	3.2	Regional Tourism Product Development and Marketing of Destinations	-	-	-	-	-	-	-	-	-
	3.3	National Tourism Marketing	68%	100%	100%	100%	99%	100%	10%	68%	7%
PA4 Local Development	4.1	Small-scale Local Investments	189%	72%	96%	75%	95%	55%	50%	94%	26%
	4.2	Inter-regional Co-operation	234%	95%	81%	73%	87%	90%	38%	114%	39%
PA5 TA	5.1	Programming, Management, Monitoring, Evaluation and Control	-	-	-	-	-	-	-	-	-
	5.2	Communication, Information and Publicity	-	-	-	-	-	-	-	-	-
	5.3	Capacity Building of OPRD Beneficiaries	80%	100%	95%	100%	100%	93%	7%	76%	5%
OP Total			144%	82%	87%	88%	94%	64%	21%	84%	12%

Source: own calculation, based on UMIS dataset, 31.12.2010

### 3.4.3 Findings of the analysis

Following the research hypotheses, the SPOT methodology provides the following conclusions:

**Figure 40 – SPOT evaluation findings by hypothesis (H3-5)**

Applied methodology	International examples	H3: objective and linked to objectives?	Objectives linked	H4: simple and robust?	H5: flexible?
Simple Progress Overview Tool (SPOT)	1 NSRF mid-term Romania	yes	Effectiveness and efficiency	yes	yes
	2 SOP-IEC mid-term Romania	yes		yes	yes
	3 OPI mid-term Malta	yes		partly	yes
	4 OPRD mid-term Bulgaria	yes		yes	yes

Source: own construction

#### **H3: Objective and linked to objectives?**

The method uses objective, quantifiable data and transforms them to facilitate the detection of issues along the implementation system. The method calculates individually meaningful ratios for further evaluation, which acts as an undoubtedly objective foundation for further exploration.

However, it is important to note that ratios need to be interpreted using reference or benchmark data. In case of process ratios (where the maximum of ratios is 100% and the minimum is 0%, following a linear function) it is objective, however ratios like popularity need other programmes or other country programmes or averages to compare to. It is time proportional values that require especially cautious interpretation, as programmes tend not to follow linear curves in terms of financial progress, but rather an exponential ones approaching the close date.

Application of the tool is also efficient and economical, as it can visualise progress (i.e. effectiveness) and also bottlenecks in the implementation system (i.e. efficiency). Its use as a dashboard makes it easy to focus only on those stages and corresponding issues in the project selection process that has outstanding values.

Therefore, the SPOT method and tool is directly linked to the evaluation theme of effectiveness, picturing how effectively (absorption) the financial objectives have been achieved. However, the “pass ratios” of the SPOT equation may act as indirect proxies to the detection of efficiency related failures in the implementation system.

#### **H4: Simple and robust?**

The method and tool is easy to comprehend, use and also easy to communicate. In fact, the power of the method partly lies in simplicity.

The method and tool is sensitive to the availability and robustness of data inputted (see International example 3 – the case of OPI Malta where robustness can only partly be identified). Inputted financial data has to be clear, available in the required structure at the same cut-off date, the financial data has to be comparable (i.e. void of exchange rate effects) and has to be available for each application stage concerned, in order to yield meaningful results.

As per robustness, the more fragmented the intervention to more uniform the distribution of grant. If the intervention covers only a few projects, than the changes occurring over time seem more dramatic (and misleading if observed at the „wrong” point in time) than an intervention with thousands of projects.

#### **H5: Flexible?**

The method and the tool is generally flexible. Thanks to the relatively extensive use of this method in the 2007-13 mid-term evaluations across EU member states, international benchmark data is available for diverse sectors, types of interventions in various countries. Based on the comparison of 4 programmes in 3 countries, there is evidence of flexibility in the method, with certain limitations.

The longer the application process, the more informative and meaningful the equation. However, when using only a few stages, the individual ratios are less meaningful (see: Malta OPI example as compared to the full data sets in the Romanian cases). The method can only be used with strong limitations in those application processes, where application stages cannot be well-defined or consistently defined.

If the evaluation covers multiple interventions, and the number of stages are different for each intervention, then a common basis has to be established as a prerequisite of comparative assessment (see Malta case and the alignment of SOP IEC and NSRF evaluations in Romania).

The SPOT method can handle financial engineering instruments with limitations. These instruments use various forms to finance interventions (e.g. loans, guarantee, venture

capital and their combinations with grants) which the beneficiaries use for their development purposes but have to pay back to the intermediary body so that it can reuse the financial resources. The actual status of such refundable sources are always the current balance of allocations. On the one hand this makes refundable instruments usable for the SPOT equation, as it shows how much resources are left in the pool available to final beneficiaries. However, this pool changes rapidly over time and the direction of change may be two-way.

The tool is also capable of acting as a monitoring guidance tool that can be integrated into a management reporting system. This can facilitate further evaluations.

### **3.5 Methodology analysis: Post contractual assessment (PCA)**

#### **3.5.1 Method overview**

The objective of the post-contractual assessment is to review (a sample of) selected projects in order to acquire a multi-faceted insight, based on qualitative information and quantitative data, into possible internal and external factors affecting progress of implementation. In order to carry out this task a sampling method which combines elements of quota sampling and judgemental sampling based on a set of characteristics and priorities have to be used. The methodology to arrive at the number of selected projects is driven mainly by the risk of possible fund de-commitment. Therefore the underlying bases for project selection are the variables which relate to the risk of losing committed grants.

This approach to the assessment of risk of fund de-commitment is based on a risk adjusted Approved (or contracted) Grant Amount (referred to as ‘Risk Severity’), and is derived as follows:

$$RISK\ SEVERITY = RISK\ IMPACT \times RISK\ FACTOR.$$

In this equation RISK IMPACT is the Approved (or contracted) Grant Amount while RISK FACTOR is a combination of two variables, namely: project status, i.e. Project progress to date; and year of approval, i.e. the year in which the n+3 rule starts. Therefore, the slower the project progress to cut-off date and the earlier the year of approval, the higher the risk factor of a project and vice versa.

A measure of risk severity (Risk adjusted Approved Grant Amount) are derived, based on each project's risk impact and risk factor. All projects can be then ranked in descending order of the total Risk adjusted Approved Grant Amount.

#### *Review of sampled projects*

This analysis examines the implementation of approved projects, with the aim of identifying any process issues which may have hindered implementation of the selected projects. This analysis is usually carried out using a twofold approach comprising of a) Planned vs. Actual expenditure and b) implementation progress.

The combined results of this analysis were also intended to shed light on the extent of possible funds de-commitment.

The planned contracting amounts and disbursements are compared with the respective actual amounts reported for each tender within each selected project. This must be done in order to highlight any possible bottlenecks.

The following table can be used to show the comparison:

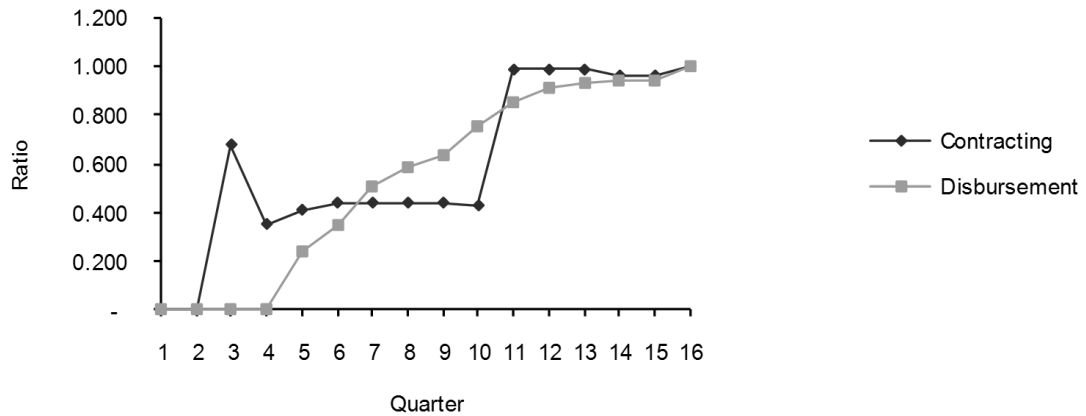
**Figure 41 – Planned vs. Actual Contracting / Disbursement Schedule in OPI, Malta evaluation PCA**

	Year Quarter	N				N+1				N+2				N+3				Total
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
Planned (€,000)	Contracting	0	260	330	1120	0	0	0	0	0	45	20	0	0	45	0	0	1820
	Disbursement	0	0	60	70	120	120	120	140	140	140	140	140	200	200	200	30	1820
Actual (€,000)	Contracting	0	0	400	200	100	50	0	0	0	0	1000	0	0	0	0	70	1820
	Disbursement	0	0	0	0	60	70	120	120	120	200	200	200	200	200	200	130	1820
N ratio	Contracting	-	-	0.678	0.351	0.409	0.439	0.439	0.439	0.439	0.427	0.986	0.986	0.986	0.962	0.962	1.000	
N ratio	Disbursement	-	-	-	-	0.240	0.351	0.510	0.587	0.636	0.758	0.848	0.916	0.928	0.937	0.944	1.000	

*Source: own construction, based on PPCD data set, 30.09.2010*

The N ratio calculated indicates the extent of the alignment of actual contracting and disbursement amounts with those planned. Therefore, an N ratio of 1 indicates perfect alignment, whereas ratios smaller than 1 would highlight delays in contracting or disbursement. Similarly, ratios greater than 1 would imply actual progress in excess of that planned. The planned versus actual contracting/disbursement schedule analysis was then used to arrive at the same analysis of the information at an aggregate level. This analysis sheds some light on the extent of the progress of contracting and disbursement vis-à-vis original planned amounts as well as a profile which may show how projects progress from year 'N' going forward.

**Figure 42 – Planned vs. Actual Expenditure in OPI, Malta evaluation, PCA**



*Source: own construction, based on PPCD data set, 30.09.2010*

For instance, as shown in the example above, project contracting starts at end of quarter 2 and achieves a 0.678 ratio by quarter 3. Delays manifest themselves in quarter 4 and continue up till quarter 10. Subsequent quarters demonstrate that the actual cumulative expenditure starts to get in sync with the planned expenditure up till project closure.

This tool enables us to highlight the quarters in which delays act as a bottleneck for the whole project; whether the project is on the right track, if still not completed, and the extent of the cumulative misalignment, if any, for each quarter.

*Planned vs. Actual Implementation schedule analysis*

This analysis is carried out to focus on the efficiency of implementation by comparing the planned implementation against actual progress for each tender within each selected project.

The data for each project is used to identify lags or leads (hereafter referred to as lag), for each project stage at tender level, thus highlighting those stages which potentially contributed to delays, and possibly hindered the smooth operation of projects, and the translation from commitments to disbursements. The resulting estimates of lag can be measured in quarters. The project stages used for this analysis can be based on the standard used in project progress report, namely Design (of tender documents), Vetting of tenders prior to publication, Tendering and contracting, Implementation and Closure.

The results were then used to arrive at an estimate of Scheduled Start Lag, Duration Lag, Stage Specific Lag and Cumulative Lag for each project stage at tender level. These lags are defined and estimated as follows:

- *Scheduled Start Lag*: This refers to the variance between the planned and actual start of the project stage.
- *Duration Lag*: This refers to the variance between the planned and actual duration of a project stage.
- *Stage Specific Lag*: This is the total lag that is attributable to a specific project stage. Therefore a lag is stage specific only when it is additional to lags brought forward from previous stages.
- *Cumulative Lag*: This lag is the total lag of a tender at each project stage. This is calculated as the sum of Stage Specific Lags at the respective project stage. Therefore, for example, the Cumulative Lag at the Tendering and contracting stage is the sum of the stage specific lag of the Design, Vetting and Tendering and contracting stages

The table below, illustrating Planned versus Actual project progress in a Gantt-type format, can be used to determine where the delay is most significant and quantify Project Lag.

**Figure 43 – Planned vs. Actual Implementation Schedule in OPI, Malta evaluation PCA**

Project Stage	Year Quarter	N				N+1				N+2				N+3				Scheduled Start Lag Quarters	Duration Lag Quarters	Cumulative Lag Quarters	Stage specific Lag Quarters
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4				
Design (D)	Planned	█																1	1	2	2
	Actual		█	█	█																
Vetting of Tender (VC)	Planned																	2	0	2	0
	Actual		█	█	█	█															
Tendering and Contracting (F)	Planned																	2	2	4	2
	Actual																	4	-3	1	-3
Implementation (I)	Planned																				
	Actual																	n/a	n/a	n/a	n/a
Closure (C)	Planned																				
	Actual																				

Source: own construction, based on PPCD data set, 30.09.2010

The preceding lead time analysis allows us to aggregate the lag figures to elicit a snapshot from a higher level. A worked example is provided on the next page:

**Figure 44 – Stage specific lags, summary of results in OPI, Malta evaluation PCA**

	Design (D) Quarters	Vetting of Tender (VC) Quarters	Tendering and Contracting (T) Quarters	Implementation (I) Quarters	Closure (C) Quarters
Project A	2	0	2	-3	
Project B	1	0	4	0	
Project C	2	2	3	2	
Project D	3	4	5	5	4
Project E	2	3	5	3	
Project F	1	1	1	0	
Mean	1.83	1.67	3.33	1.17	
Mode	2.00	0.00	5.00	0.00	
Median	2.00	1.50	3.50	1.00	
Standard Deviation	0.75	1.63	1.63	2.79	

Source: own construction, based on PPCD data set, 30.09.2010

The table above helps us to highlight two important aspects: at project level, we can immediately identify which projects are at risk of fund de-commitment. Significant delays in all project stages (example Project D) would signal that part of the committed funds would not be eligible for reimbursement if the project fails to be completed on time. The summation of the Approved Grant Amount of those risky projects would give an indication of the total Euro amount which is at risk. At stage level, we can identify where attention needs to be directed with a view of eliminating inefficiencies (and hence improving timely effectiveness) and compressing time lags within specific Project Stages.

### 3.5.2 International examples

#### International example 1 and 2 – OPI and OPII mid-term evaluation, Malta

The mid-term evaluation of the OPI and OPII (NSRF) in Malta was commenced by PPCD in 2010 and concluded in 2011. Cut-off date for the evaluation was 30 April, 2010. As described previously, Malta was in a unique situation in many aspects, out of which it is of major importance that the Maltese NSRF was a Single Programming Document (SPD) consisting of two Operational Programmes, i.e. OPI (containing ERDF) and OPII (containing ESF). The mid-term evaluation of both OPs took advantage of the application of the PCA methodology. As a first step of the assessment, a sample was created with regard to the following parameters:

- Fund used (i.e. ERDF, ESF, CF, based on approved grant distribution at cut-off)
- OP (i.e. OPI and OPII, based on approved grant distribution at cut-off date)
- Priority Axis (PA1-5, based on approved grant distribution at cut-off date)
- Project status (i.e. pre-defined relevant progress stage distribution at cut-off date)

**Figure 45 – Summary of stage specific lag estimates in OPI, Malta evaluation PCA**

	Design (D) Lag Quarters	Vetting of Tender lag (V) Quarters	Tendering and Contracting lag (T) Quarters	Implementation lag (I) Quarters	Closure lag (C) Quarters	Cumulative lag Quarters	Planned project duration Quarters	Total Planned Contracted Amount €
Average	0.61	0.59	0.43	0.08	(0.01)	1.46	15	
<b>Total planned contracted</b>								<b>267,542,440</b>
<b>Total planned contracted above threshold</b>								<b>24,128,239</b>

Source: own construction, based on PPCD data set, 30.09.2010

**Figure 46 – Summary of project specific lag estimates in OPI, Malta evaluation PCA**

Project No	Design (D) Lag Quarters	Vetting of Tender (V) Quarters	Tendering and Contracting (T) Quarters	Implementation (I) Quarters	Closure (C) Quarters	Cumulative lag Quarters	Planned project duration Quarters	Total Planned Contracted Amount €	Cum Lag / Duration %
ERDF-058	-	0.33	-	0.67	1.00	2.00	8.00	2,569,064	25.0%
CF - 123	0.40	-	-	(0.40)	0.60	0.60	16.00	7,440,000	3.8%
ERDF-017	1.38	0.38	-	(0.17)	1.00	1.77	13.00	17,355,905	13.6%
ERDF-135	1.00	-	-	0.60	(0.67)	1.00	26.00	10,000,000	3.8%
ERDF-048	(1.00)	1.00	1.25	0.75	(2.00)	1.00	14.00	16,955,825	7.1%
CF - 124	1.78	0.13	0.67	0.67	-	1.91	16.00	13,463,000	11.9%
ERDF-133	(2.67)	-	(1.00)	-	N/A	(3.67)	20.00	10,000,000	-18.3%
ERDF-068	(1.25)	2.75	1.00	0.33	N/A	2.75	11.00	2,742,635	25.0%
ERDF-033	1.22	0.70	0.38	(2.60)	0.50	0.90	16.00	14,289,174	5.6%
ERDF-104	0.17	0.36	0.50	(0.25)	-	0.67	13.00	9,291,411	5.1%
ERDF-047	-	1.33	1.00	(1.00)	N/A	1.25	13.00	7,107,315	9.6%
ERDF-012	1.86	0.43	0.86	1.00	N/A	3.71	8.00	5,011,998	46.4%
ERDF-031	1.80	1.67	(0.67)	-	(1.00)	2.20	19.00	5,173,091	11.6%
CF - 117	0.14	-	0.40	1.00	-	0.63	16.00	51,624,713	3.9%
ERDF-142	1.17	-	0.90	-	N/A	1.19	15.00	1,018,910	7.9%
ERDF-032	(0.23)	1.00	0.83	-	-	1.00	21.00	8,679,202	4.8%
ERDF-072	0.20	1.00	(0.20)	0.50	N/A	1.40	10.00	1,420,897	14.0%
ERDF-081	1.25	0.65	1.45	0.68	N/A	4.00	10.00	4,231,445	40.0%
ERDF-199	0.33	0.67	(0.33)	(0.67)	0.50	0.20	10.00	12,140,095	2.0%
ERDF-088	N/A	N/A	N/A	N/A	N/A	N/A	16.00	8,851,227	
ERDF-001	1.29	0.48	0.07	0.70	N/A	2.67	16.00	15,237,500	16.7%
ERDF-054	4.00	-	3.50	N/A	N/A	5.50	7.00	1,613,409	78.6%
ERDF-134	(1.67)	-	(0.67)	-	N/A	(2.33)	21.00	2,500,000	-11.1%
ERDF-039	0.37	0.57	0.44	(0.21)	N/A	0.85	28.00	30,865,936	3.0%
ERDF-064	3.20	0.60	(0.13)	0.29	-	3.90	14.00	7,959,688	27.9%

Source: own construction, based on PPCD dataset, 30.09.2010

The above results clearly show that the delay was fundamentally attributable to design stage and more specifically, as a duration lag. This means, that the absorption risk carried by the projects in the representative pool (the sample) were the result of late start of implementation design (minor factor) and the long duration in implementation design tasks (major factor).

As further assessment disclosed, these two reasons were mainly related to competency and capability issues in terms of public procurement knowledge. These issues entailed law disputes in several cases that prolonged the tendering and contracting processes significantly.

### 3.5.3 Findings of the analysis

Following the research hypotheses, the PCA methodology provides the following conclusions:

**Figure 47 – PCA evaluation findings by hypothesis (H3-5)**

Applied methodology	International examples	H3: objective and linked to objectives?	Objectives linked	H4: simple and robust?	H5: flexible?
Post-Contractual Assessment (PCA)	1 OPI mid-term Malta	yes	Accountability (indirectly efficiency)	partly	yes
	2 OPII mid-term Malta	yes		partly	yes

*Source: own construction*

#### **H3: Objective and linked to objectives?**

The overall advantage of this method is that it provides a high level, objective insight into the reasons behind the crucial factors of project implementation (and absorption) risks. The method has a lot of advantages and a few limitations, which makes it a suitable tool for setting the scenes for the evaluation of aggregated absorption risk in grant programmes.

This analysis is not capable of disclosing the reasons behind the delay. These reasons might be investigated through interviews or direct surveys with stakeholders, i.e. project owners, their suppliers and relevant authorities. However, the outputs of the method may triangulate and focus further explorations of the reasons for the lags, enhancing the efficiency of the process.

The PCA method mainly and directly raises questions on the efficiency of the operations of the institutional system and indirectly detects issues of accountability.

As per efficiency, the method points at those sub processes in the implementation process, where lag is generated that is rolled on to latter sub processes. This might have a consequence on absorption (observing the  $n+2/n+3$  rule), capacity (peak time management of personnel at IBs and MAs) and project timing (e.g. construction limitations at wintertime).

Accountability issues are detected indirectly as lags mean that certain regulatory deadlines have not been met, or rules such as  $n+2/n+3$  or public procurement timing requirements have not been observed carefully.

#### **H4: Simple and robust?**

The methodology is easy to build up and easy to communicate, which facilitates the potential acceptance of both the approach and the results derived via its use.

The application of the methodology addresses two simple questions that policy makers and programmers are most eager to learn. First, it answers the question of “How much is at risk?”, which is described through the extent of variance between planned and actual contracting/disbursement amounts. Second, the methodology addresses the question of “Where does the risk come from?”, showing which tendering or project stage is prevalent in the delay.

The tools are sensitive to data clarity and robustness issues in both the sampling stage and the analysis stage. Therefore in case of doubt, a data clarification process has to be initiated prior to inputting in the model in order to avoid false or misleading results. Also, considerations have to be made prior to the analysis to minimise selection bias in the sampling process.

The data required for the evaluation is compulsory by EU regulations to be available to the implementation institutional system. Also, only a few obvious datum is needed per projects, which makes them quick to collect and input.

#### **H5: Flexible?**

The tool is flexible and can be tailored to the needs and context. This constitutes both an advantage (method can be tailored to data structures and data quality) and a risk (selection bias and data clarity issues) as described previously.

Also, the sampling element carries numerous risks, therefore this exercise has to be carried out with regard to the most characteristic stratifying features, such as programme or priority, territorial, sectorial and beneficiary type split.

The methodology and its tools can be integrated into a monitoring and reporting system so that an early warning system or a sound basis for an on-going evaluation can be developed.

## **3.6 Methodology analysis: Lead time analysis (LTA)**

### **3.6.1 Method overview**

Lead time analysis is a simple but powerful tool for measuring an important aspect of efficiency in a grant system. Lead time analysis is used to objectively judge how much time is spent in between the relevant stages of the entire application and implementation process. Relevant in this context is a result of pre-definition: it depends on the scope and depth of the analysis. Most often, however, the subject of the assessment is either the application process, or the implementation process, or both. As a first step to starting the assessment, the relevant stages and corresponding data has to be defined (e.g. which stages are required by the analysis).

The potential outputs of a lead time analysis are twofold: first, the assessment may shed light on where the bottlenecks in the process are, and what their severity is (internal assessment) and second, output figures can be compared across the institutional system or other benchmark data in order to identify which entities require more time for a certain process and what is the reason for that (e.g. capacity shortage, capability or competence issues).

#### **Project selection process**

As a first step to understanding the context, the grant allocation process (e.g. an application process or other type of project selection and implementation process) has to be clearly defined. In order to demonstrate the use of the lead time assessment method, a standard grant application process will be used, consisting of a three-level project selection mechanism including an administrative, an eligibility and a technical assessment on the application submitted. It is presumed that if an application passes through all three levels, then it is approved, a grant contract is signed, then implementation and corresponding payment can be initiated to the beneficiary.

A standard application process for demonstrating purposes includes seven major steps from submission of application up to payment of the grant (see Figure below). Please note that in some of the member states the three-stage (administrative, eligibility and technical) application assessment process is reduced to two (see the Bulgarian example later in this article) or even one in practice by integrating one stage into another.

**Figure 48 – A standard project selection (application) process using a three-stage selection system**



Source: own construction

**Method**

The purpose of the lead time assessment is to create a sound and objective basis for the evaluation of efficiency of the grant allocation and implementation process. In this context, lead times can be defined as time spent between any two relevant stages of the grant process. This objective information then can be used to compare with other figures and to explore reasons behind the facts.

The execution of the method requires two steps: 1) data collection and verification and 2) analysis. Data collection regards project application data available in the monitoring and information system. It is important to highlight, that lead time analysis may use project samples, and however, it is a method that prefers robust and complete data set to be available for the assessment. Data should be verified by the assessment of the monitoring and information system (data structures, methodologies used such as data conversion methods, and data completeness), which is to be coupled with a small sample for pre-testing (pilot). The table below shows an example of a data input table required for the assessment:

**Figure 49 – Lead time analysis standard data collection sheet**

		Project submission	Admin assessment	Eligibility assessment	Technical assessment	Approval	Contracting
Lead time from “Project submission” (days)	Min	Day 0					
	Max						
	Mean						

Source: own construction

The LTA might concentrate on a number of indicators for each project, at least on:

- *Application assessment time*: days spent from submission of the project application up till technical assessment
- *Application approval time*: days spent from technical assessment to contracting
- *Contracting time*: from submission to contracting (i.e. the sum of the previous two indicators)

The above figures for each projects are to be aggregated by the organisation responsible for the processing of project applications and supporting and monitoring project implementation tasks. The absolute values of these aggregates enable assessment on three levels: 1) assessment of the figures themselves, 2) comparison with other organisation of the fund management institutional system, and 3) comparison with similar organisations in other member states.

The next two examples show how in Bulgaria (mid-term evaluation of the 2007-13 Regional Operational Programme) and Malta (mid-term evaluation of the 2007-13 European Social Fund based OPII Operational Programme) used the lead time assessment and what results its application could bring.

### 3.6.2 International examples

#### **International example 1 – OPRD mid-term evaluation, Bulgaria**

The lead time analysis of the OPRD focused on the project selection process from submission of the application, through technical (also including admin and financial) check to approval and then contracting. These stages represented the relevant points of the assessment. The figures drawn as a result set the scene for the evaluation of reasons behind the durations that were longer than expected or adequate. Part of the assessment concentrated on the identification of bottlenecks in the process, and also the comparison of figures with national regulations on provisions governing the deadlines for the implementation system.

The assessment found that the length of the total lead times largely depended on project selection methods. Schemes with a rolling submission procedure (where projects submitted are evaluated in batches and are accepted until financial resources run out) had longer average total lead time (127 days), while projects with a fixed deadline (where projects submitted are prioritised during evaluation and above a certain score all receive the grant) had a slightly shorter total lead time (120 days). The total lead time from registration of applications to contract signing took 118 days for OPRD projects, according to the analysis. The Figure below shows the operations of the OPRD along with the relevant lead time figures between the identified stages of the application process.

**Figure 50 – Summary table of lead times in OPRD, Bulgaria evaluation**

Operation	Contracted projects	From registration to technical and financial check	From technical and financial check to approval	Total lead time: from registration to contracting
1.1	133	125	10	139
1.4	70	70	11	120
2.1	58	93	16	120
3.1	10	51	14	142
3.3	9	38	7	61
4.1	108	74	15	111
4.2	55	71	19	125
5 (TA total)	43	38	16	58
<b>Total</b>	<b>486</b>	<b>89</b>	<b>13</b>	<b>118</b>

*Source: own calculation and construction (based on UMIS data set, 31.12.2010)*

As the table above shows the applications of Operation 1.1 and Operation 3.1 had to go through the longest process, with Operation 1.1 being the most popular Operation, with a logical reason that the high number of applications caused heavy workload for the MA to process. The TA projects (Operation 5) have been processed very quickly which reflected the special character of TA. Operation 3.3 was also very quick, thanks to the low number of projects required to be processed. Placing the overall lead time result to an international context highlights the fact that the Bulgarian OPRD total lead time result at the end of 2010 was somewhere in between the corresponding value of the Romanian Regional OP (292 days) days and the seven Hungarian Regional OPs' average of 63 days.

*From registration to technical and financial check*

**Figure 51 – Lead time analysis from registration to technical and financial check OPRD, Bulgaria**

Operation	Average	Min	Max	Deviance
1.1	125	51	262	42,3
1.4	70	1	126	29,8
2.1	93	13	271	48,3
3.1	51	36	99	22,2
3.3	38	27	43	8,0
4.1	74	32	124	17,2
4.2	71	71	80	1,0
5 (TA total)	38	1	141	32,6
<b>Total</b>	<b>89</b>	<b>1</b>	<b>271</b>	<b>43,3</b>

*Source: own calculation and construction (based on UMIS data set, 31.12.2010)*

The average lead time from registration to the technical and financial check was 89 days. In the case of Operation 1.1 this process took significantly longer (125 days), while in Operation 3.3 shorter, 38 days, respectively. The deviation of lead times of Operation 2.1 applications was the most spectacular: it varied from 13 to 271 days to reach the technical and financial check from the registration. The case of Operation 1.1 was also similar, with lead times ranging from 51 to 262 days. The next Figure shows the difference of pattern when separating applications of the OPRD with rolling submission and set deadline procedures.

**Figure 52 – Lead time analysis rolling and set deadline procedures OPRD, Bulgaria evaluation**

Rolling submission procedures					Set deadlines procedures				
Operations	Average	Min	Max	Deviance	Operations	Average	Min	Max	Deviance
1.1	140	65	208	28,4	1.1	84	51	262	48,3
1.4	-	-	-	-	1.4	72	25	126	29,0
2.1	100	27	168	50,1	2.1	-	-	-	-
3.1	-	-	-	-	3.1	49	36	87	21,5
3.3	-	-	-	-	3.3	-	-	-	-
4.1	75	32	108	18,5	4.1	73	46	124	14,9
4.2	-	-	-	-	4.2	71	71	80	1,0
5 (TA total)	-	-	-	-	5 (TA total)	-	-	-	-
<b>Total</b>	<b>112</b>	<b>27</b>	<b>208</b>	<b>42,3</b>	<b>Total</b>	<b>71</b>	<b>25</b>	<b>262</b>	<b>29,3</b>

Source: own calculation and construction (based on UMIS data set, 31.12.2010)

In the figure above the first table shows the rolling submission procedure, where no deadline had been set and the second one shows the procedures with a fixed deadline. There were huge differences between the two project selection methods. At rolling selection procedure the average lead time was significantly higher (112 days), than the procedures with a certain deadline, where the average lead time was 71 days, which was low compared to the overall average lead time from registration to technical check (89 days). At the set deadline procedures the deviance has also decreased to 29.3 days. So in overall, the project selection with a set deadline led to a significantly lower lead time (from registration to technical and financial check) and also to lower deviance.

*From technical and financial check to approval*

The other relevant assessment period of the application system is the time spent between a positive recommendation on the provision of grant to a beneficiary by the intermediary body and the official decision of the management authority. This is a rather administrative action indicating the smoothness of cooperation between the actors of the management and implementation system.

With an average of 13 days the process from technical and financial check to approval took the shortest time among the lead times in between the relevant stages of the application process. The average varied between 7 (Operation 3.3) to 19 (Operation 4.2) days to approve applications.

**Figure 53 – Lead time analysis from technical and financial check to approval OPRD, Bulgaria**

Operation	Average	Min	Max	Deviance
1.1	10	2	17	5,3
1.4	11	1	18	4,3
2.1	16	2	33	11,8
3.1	14	10	33	3,2
3.3	7	5	10	2,5
4.1	15	8	26	5,7
4.2	19	19	19	0,0
5 (TA total)	16	1	39	15,4
<b>Total</b>	<b>13</b>	<b>1</b>	<b>39</b>	<b>7,3</b>

Source: own calculation and construction (based on UMIS data set, 31.12.2010)

*Total lead time: from registration to contracting*

A summary of the individual results of the lead times in between single statuses showed that the average total lead time was 118 days with OPRD applications. Operation 3.1 led the rank in terms of average lead time with a figure of 142 days. Operation 3.3 had the shortest average time (except TA) that passed from the registration of application to contracting which took 51 days. The shortest individual lead time belonged to Operation 1.4, with only 1 day.

**Figure 54 – Lead time analysis from registration to contracting results OPRD, Bulgaria evaluation**

Operation	Average	Min	Max	Deviance
1.1	139	81	283	42,8
1.4	120	1	211	53,4
2.1	120	30	481	62,8
3.1	142	97	149	16,1
3.3	61	51	81	15,0
4.1	111	62	548	64,4
4.2	125	110	428	62,9
5 (TA total)	58	2	182	41,8
<b>Total</b>	<b>118</b>	<b>1</b>	<b>548</b>	<b>58,4</b>

Source: own calculation and construction (based on UMIS data set, 31.12.2010)

## **International example 2 – OPII mid-term evaluation, Malta**

The evaluation used two lead time analyses, one for the application process and another for the invoice payment process for OPII, which was the 2007-13 Operational Programme governing all European Social Fund related interventions in Malta.

For easy comparison with Bulgarian figures, only the application process lead time assessment is presented in this section. Please note that the Bulgarian case counts in standard days, while the Maltese example contains a more advanced view by calculating in working days that enables capacity comparisons in terms of the management and implementation institution system.

The analysis was carried out in order to provide an insight into the efficiency of the application process at the Managing Authority and the Intermediary Body, in respect of OP II related activities. The data included a series of tracking dates spanning the entire application process, as applicable, for the selected and non-selected applications.

The tracking dates included: application submission date, eligibility check date, evaluation result date, plus for selected projects (letter of offer date, Grant Agreement date), and for non-selected applications the letter of rejection/non-selection date.

The data was then analysed to arrive at the number of working days (taking into consideration the local calendar of national and public holidays) elapsed between each pair of sequential events as described above and the cumulative days elapsed from submission of application to three events as follows: from submission to letter of offer (for selected applications), from submission to Grant Agreement (for selected applications) and from submission to letter of rejects/non-selection (for non-selected applications).

It is important to note that applications that were still under evaluation as at cut-off date were not included as the detailed tracking data was not available and event pairs completed on the same date were treated as equivalent to 0 working days.

The lead time analysis shown in the figures on the next page reflect the analysis in respect of the application process at the MA and the relevant IB, by the respective call for applications and/or aid scheme as applicable.

The entire process in the case of Malta took an average of 193.7 days (Operation level means min: 148 days, max 245.3 days) for the Managing Authority to accomplish on average in between submission of an application and the letter of offer stage (very similar to the approval stage in the Bulgarian context). Deviances were great in between the Operations. As the Figure shows, it was the submission to eligibility check stage that was clearly responsible for at least half of the lead time duration in the process with an average value of 116.6 days (Operation level means ranging from min 68.1 days to max: 157.8 days). However, in the case of interventions, where it was the Intermediary Body processing the applications (see next Figure for reference), the process was much shorter: 19.9 days on average (min: 14.8, max: 25 days).

For a quick comparison with Bulgarian figures, there are many aspects that makes this OP different: there are fewer number of projects, the character of purely ESF interventions of the Maltese OPII is different from the generally ERDF based regional development interventions of the Bulgarian OPRD. However, the figures show an almost double value in the case of Malta which also underpinned the necessity to evaluate the reasons behind the facts.

**Figure 55 – Lead time analysis of the application process by calls OPII, Malta evaluation**

Managing Authority								
	No. of applications	Submission to Eligibility Check	Eligibility Check to Evaluation Result	Evaluation Result to Letter of Offer	Letter of Offer to Grant Agreement	Submission to Letter of Offer	Submission to Grant Agreement	Submission to Rejection letter
ESF (By Call)								
Call 1 - OP II - ESF	16	68.1	56.0	72.5	251.7	187.3	412.7	112.0
Call 2 - OP II - ESF	5	115.6	33.0	9.0	No data	148.0	No data	141.5
Call 3 - OP II - ESF	37	102.2	28.3	50.0	277.5	169.3	449.7	137.4
Call 4 - OP II - ESF	6	125.7	45.7	56.7	No data	228.0	No data	168.0
Call 5 - OP II - ESF	28	157.8	54.4	36.0	122.8	245.3	327.0	178.4
Total No. of Applications	92							
<b>Overall Average</b>		<b>115.5</b>	<b>39.2</b>	<b>46.0</b>	<b>230.2</b>	<b>193.7</b>	<b>409.4</b>	<b>147.0</b>

Source: own calculation, based on PPCD dataset, 30.09.2010

**Figure 56 – Lead time analysis of the application process by schemes (MA and IB) OPII, Malta**

OP II - ESF Lead time analysis of the application process								
Intermediate Body								
Aid Scheme	No. of applications	Submission to Eligibility Check	Eligibility Check to Evaluation Result	Evaluation Result to Letter of Offer	Letter of Offer to Grant Agreement	Submission to Letter of Offer	Submission to Grant Agreement	Submission to Rejection letter
EAP	1,047.0	12.8	-	4.1	5.5	14.8	19.6	16.1
TAF	1,084.0	22.3	32.6	8.8	12.5	25.0	51.5	65.2
Total No. of Appl.	2,131.0							
<b>Overall average</b>		<b>17.3</b>	<b>16.6</b>	<b>5.7</b>	<b>8.2</b>	<b>19.9</b>	<b>35.9</b>	<b>40.8</b>

Source: own calculation, based on PPCD dataset, 30.09.2010

### 3.6.3 Findings of the analysis

Following the research hypotheses, the LTA methodology provides the following conclusions:

**Figure 57 – LTA evaluation findings by hypothesis (H3-5)**

Applied methodology	International examples	H3: objective and linked to objectives?	Objectives linked	H4: simple and robust?	H5: flexible?
Lead time assessment (LTA)	1 OPRD mid-term Bulgaria	yes	Efficiency (indirectly accountability)	yes	yes
	2 OPII mid-term Malta	yes		yes	yes

*Source: own construction*

#### **H3: Objective and linked to objectives?**

The method is objective, promotes the establishment of a solid, evidence-based foundation for the further use of qualitative techniques (such as interviews of consultations for revealing the potential factors behind the figures calculated).

The outputs of lead times assessment are figures, days which can be attributed to processes or their parts, organisations and other capacities.

It is a common evaluation mistake that this method is only used to provide a stock indicator at the cut-off date for the evaluations („now”). In order to learn more about the learning capacity and adoptability of the implementation system, a dynamic assessment („up till now”) can also be conducted.

Lead time analysis enables objective assessment of an important facet of implementation system efficiency, while also pointing at certain issues in terms of accountability. Lead times generally vary based on the timing of the assessment. When starting a programme, the learning curve effect hinders operations on both the beneficiary and implementation system side. In later programme stages figures tend to converge to the real values, assuming that call portfolio and call conditions are rather stable (assuming that no major change has occurred in the call structure or processes).

#### **H4: Simple and robust?**

Lead time figures calculated are easy to communicate, however they are to be interpreted with caution. The characteristic of the application handled by organisation may be significantly different. For instance, organisations handling small-scale mass applications

for small and medium enterprise equipment procurements will presumably have lower lead times than those organisations which manage only a few but large infrastructure development projects. Therefore programme features such as call complexity, potential project size, preparatory and regulatory requirements, potential number of applicants have to be considered along with the figures when making comparisons and setting up hypotheses.

Not only the absolute values and the means, but also the deviation of lead time figures carry important information for evaluation. Considerable deviation reflects the unbalanced operation of the organisation that might be the result of inadequate call definition, issues with scheduling or peak time capacity management or competence related considerations.

It is important to note that when formulating evaluation findings, there is always at least three stakeholders to attribute the objective figures to: 1) the regulator 2) the implementing bodies (MA and IB(s), and 3) the applicants. However, please note that issues on the applicant side should usually be handled through interventions in the operations of the regulator and the implementing bodies.

#### **H5: Flexible?**

The LTA method is flexible in the sense that the scope can be tailored to the length and specifics of the implementation system (e.g. number of processing stages) and the requirements and scope of the evaluation. In order to enable international comparisons, the stages of the application pipeline under assessment have to be in line with those to be compared to.

The method uses monitoring data that is compulsory for each member state to collect. Therefore it is unlikely that the assessment cannot be carried out properly as a result of lack of information or ability to access data. However, data clarity and completeness issues might arise.

The figures calculated can be further used for other analyses, for instance to project related absorption risk assessment (if a project approaches  $n+2$  deadlines).

The calculation of figures can be easily automatized and integrated into a monitoring system for providing real-time reporting.

## 4 CONCLUSIONS

This section contains the conclusions of my research conducted following my five hypotheses tested through two analyses. Section 4.1 contains the answers to the first hypothesis (H1) on the results of Cohesion Policy in the EU-12 regional development, Section 4.2 summarises my conclusions on my second hypothesis (H2) on EU-12 ROP diversity and uniformity in 2007-13 while Section 4.3-4.5 cover the answers to the last three hypotheses (H3-5), all focusing on evaluation methodology.

All methodologies described and analysed in this study (see Sections 3.4, 3.5. and 3.6) have been developed, tested, tailored and simplified through a minimum of 5-6 years in the course of iterative processes. All of them have been applied in several member states and have been accepted by local authorities and the EU Commission as methods capable of addressing and answering their evaluation questions. In conclusion, each of these methods provide a safe and cost-effective basis for evaluation, with regard to their potential advantages and limitations.

The table below summarises the main findings of my research on H3-5 through providing a brief answer to these hypotheses in relation to the methodology used. This table provides an overview of the aggregated results that will be further detailed in the next sub-sections.

**Figure 58 – Summary table of conclusions by hypothesis (H3-5)**

Applied methodology	International examples	H3: objective, linked to objectives?	Objectives linked	H4: simple and robust?	H5: flexibility?
<b>Simple Progress Overview Tool (SPOT)</b>	1 NSRF mid-term Romania 2 SOP-IEC mid-term Romania 3 OPI mid-term Malta 4 OPRD mid-term Bulgaria	yes	Effectiveness and efficiency	yes	yes
<b>Post-contractual assessment (PCA)</b>	1 OPI mid-term Malta 2 OPII mid-term Malta	yes	Efficiency (indirectly accountability)	partly	yes
<b>Lead time analysis (LTA)</b>	1 OPRD mid-term Bulgaria 2 OPII mid-term Malta	yes	Efficiency (indirectly accountability)	yes	yes

*Source: own research conclusions*

Having assessed ROPs of the EU-12 countries and the regional results of Cohesion Policy interventions and the content, application and result of the preselected prominent methods of the 2007-13 EU mid-term programme evaluations, my conclusions by hypothesis are as follows:

## 4.1 Thesis 1 – Cohesion Policy regional results

**Hypothesis 1: EU Cohesion Policy could contribute to the development of regions in the EU-12, however, it could not counteract the growing regional disparities within these countries.**

The objective of EU Cohesion Policy is the moderation of regional inequalities. Though Cohesion Policy has justifiably supported development in the EU-12 countries in the 2004-2011 period (EC, 2014d), it has failed at mitigating sufficiently the increasing interregional polarisation within the EU-12 countries according to my analysis.

Apart from the most widely acknowledged development indicator, GDP per capita, I based my assessment on two relevant items of regional development theory: the *Williamson hypothesis* describing a relation between development and homogeneity that resembles a 'U' upside down (Williamson, 1969), and the neoclassical theory based *beta convergence* phenomenon (Canova, 2004; Dvoroková, 2014) on the diminishing marginal returns on capital, claiming that underdeveloped regions grow at a rate higher than developed regions.

The diversity is obvious in trends and results across the EU-12. The beta convergence effect is underpinned by the relative development figures of the EU-12 since the accession times. At the regional level, however, interregional inequalities in countries have been disclosed through the provision of the minimum and maximum relative GDP/capita figures of the NUTS2 regions.

The changes in the 2004-2011 period in the countries of the EU-12 with more than one NUTS2 level territorial unit are indeed dramatic: there is an overall improvement effect assumedly partly deriving from the EU co-financed interventions - similarly to other previous study findings - however, the development gaps within countries has grown enormously between 2004 and 2011. In some cases, relative GDP/capita figures compared to EU-28 average figures have more than doubled between 2004 and 2011.

More developed countries suffered less from increasing gaps, while relatively underdeveloped ones has shown outstandingly growing gap figures. This phenomenon underpins the Williamson hypotheses, as in relatively more developed countries (the Czech Republic, Hungary, Slovenia), the development gap between regions did not increase in proportion with the huge leap of figures in the less developed countries (Bulgaria, Romania, Poland, Slovakia).

According to the last, accessible self-assessment on the performance of Cohesion Policy by the European Commission (EC, 2014), the Cohesion Policy contributed to the annual GDP growth of 0.5-1.0% in the EU-12 countries in the 2007-13 period, providing evidences on its contribution to member state development. However, the report does not formulate straightforward statements on the changes of regional inequalities, though it is a fundamental role of the Policy.

As a conclusion, the attributability of both the beta convergence phenomenon and the Williamson hypothesis could have been identified in the topic of development of EU-12 countries and regions in the 2004-2011 period, indicating that Cohesion Policy has indeed supported development in these countries, but at the same time it has failed in sufficiently mitigating the increasing regional polarisation within the EU-12 countries.

## **4.2 Thesis 2 – ROP diversity and similarity**

**Hypothesis 2: Regional Operational Programmes in the EU-12 countries in 2007-13 are diverse in terms of their approach and their relative weight in the NSRF despite their content (interventions) and their mid-term evaluations are very similar.**

The analysis identified three distinct approaches used in the selected NSRFs in the EU-12 in the 2007-13 EU budgetary period. Member states either chose to have a ROP for each region, or one for the entire country or to position ROP measures within another programme, the choice of which was logically dependent on factors such as size (i.e. territory), uniformity and national level regional development attitude of the country. 2007-13 approaches live on in the 2014-20 period as well, though there have been countries shifting from one approach to another (e.g. Hungary used to have 7 ROPs in the 2007-13 period while in 2014-20 it has one for the less developed regions together and one for the Central Hungarian Region (i.e. the compulsory minimum). Should there be no EC requirement to separate regional interventions based on the relative development indicators of a region, Hungary is likely to have only one ROP.

The analysis also disclosed a high concentration of ROP measures in the funding structure, with the four top countries accounting for 84% of such funds, while 8 countries share the remaining 16%. The relative weight of the ROP measures in the respective NSRFs shows a great spread, ranging from 5% to 34%. The reason for this lies partly in development approach as well: member states in 2007-13 and previous periods had to decide how to balance sectoral approach with territorial approach. Note that this has

slightly changed in 2014-20 with the introduction of thematic objectives in CPR (EC, 2014a) and ERDF regulations (EC, 2014b).

Another factor behind the allocations is the relative weight of topics that traditionally belong better to the territorial aspect, i.e. tourism, industrial parks, regional level accessibility, transport, local social, educational and economic issues. These interventions are very similar in all EU-12 regional cases, with slight differences in emphasis, assumedly based on relative development status and institutional system and authority.

It is an obvious disadvantage at the choice of interventions that they target operational level objectives (means), but only indirectly pursue strategic level goals like extended regionalism (Horváth, 2015), more authority to regions or strengthening the second level of cities, i.e. (potential) regional centres to ease the polarised situation governed by the overwhelming power of the centre, with unmatched development differences compared to the countryside.

The ROP mid-term evaluations chosen for the research have shown clearly that the selected evaluations did not follow the focus areas and intensities pre-set by the European Commission (EC). The focus of the evaluations was more intense than the EC had outlined. Evaluation of rationale and impacts were moved to the back, while relevance, effectiveness and efficiency were all highlighted topics. In fact, three out of the five evaluation aspects were differently handled. I have to note, however, that these changes are not dramatic: there is only one step forward or backward on the four-level scale (from 0 to 3 stars, in intensity). What is more dramatic, is the narrowing of the scope of the evaluation from five to three topics, as two of them have been practically ignored in the course of the evaluation (in line with the expectations of the commissioners of the evaluations).

The concentration of topics showed high interest of the programme implementers in topics that are closely linked to *implementation* (effectiveness and efficiency) as opposed to *planning* (rationale, relevance). However, it is important to note that the general purpose of mid-term evaluations is the right mixture of the two: in the mid-term phase, there are already results at hand that enable the drawing of conclusions and initiating modifications to enhance performance in the currently ongoing programme. On the other hand, however, mid-term evaluations are commissioned in time periods, when

programming of the next programme begins. As a result, conclusions drawn in mid-term are immediately integrated in the planning of the next programme.

In the course of the implementation, it is understandable that implementers, politicians, and the broader public are more interested in current status and spending perspectives. There is a competition for progress among countries and within countries to absorb funds. The higher emphasis of efficiency in the analysis carries two messages in my interpretation: first, in the countries that I selected there is a long track record of keeping an eye on accountability, and second, it was a clearly articulated need of these member states to reduce spending on the delivery of the funds (i.e. transaction costs) and improve the ratio of result / cost by increasing results, lowering costs or both.

However, the fact that efficiency caught up to effectiveness (claimed to be the single strong focus of mid-term evaluations by the EC) indicates that selected member states were much more keen to learn *how* they perform than *what* they perform. Losing the focus of achieving pre-set objectives and concentrating more on fine-tuning delivery is not entirely in line with the original intentions of the policy makers and the programmers.

The hypothesis has been justified, ROP measures in the EU-12 countries in 2007-13 are indeed diverse in terms of their approach and their relative weight in the NSRF despite their content and their mid-term evaluations is very similar.

### **4.3 Thesis 3 – Method objectivity, and link to intervention objectives**

**Hypothesis 3: The selected evaluation methodologies are capable of providing objective<sup>17</sup> basis for evaluation and address at least one of the general intervention objectives of effectiveness (and relevance), efficiency and accountability<sup>18</sup>.**

The SPOT, PCA and LTA methodologies detailed in this study are capable of establishing an objective foundation and setting the scene for further evaluation. This evidence-based, qualitative basis provides opportunity for exploration of reasons behind which fall into the category of subjective and formative evaluation methods (e.g. consultations, surveys,

<sup>17</sup> For the notion of objectivity, I integrate and use the notions „factual” and „unbiased” (Scriven, 1972) and „that disregards expertise of biased persons” (Scriven, 1997) and „quantitative” (Terstiege, 2012). I do not consider „distance” as I agree with (Scriven 1972) that „Distance does not guarantee objectivity, it merely guarantees distance”

<sup>18</sup> see MEANS, 2000

interviews). Further investigation of reasons behind the facts may alter or underpin the initial finding based solely on figures. Using soft drilling techniques to interpret the objective results is not an option but a requirement to yield real benefits. The objective, quantitative approach preceding the subjective, qualitative techniques provides a logical framework for the evaluations, in line with the community guidance of the triangulation principle.

First, it is more logical and cost-efficient to start consultations having the facts and evidences at hand. Management of an evaluation is a multi-stakeholder process with lots of iterations. However, the number of iterations can be reduced drastically if there is a common, objective understanding among stakeholders, which can be established through the results of a quantitative analysis. Therefore, an evidence-based, objective approach helps to build trust and acceptance.

Second, consultations can be more targeted, more effective if there are certain focus points of the investigation, which can be established via quantitative, objective methods. This narrowing of the scope of further explanation not only facilitates the management of stakeholders, but also assists evaluators to concentrate their resources to areas which require a more in-depth analysis based on findings of the objective analysis.

Third, subjective elements, such as consultations may be used to challenge, cross-check and validate objective findings. In case of mismatch, the reasons must be observed on both sides (an inconsistency can obviously derive from interests or opinions, however, data clarity and input failures also lead to improper results). In such a case an iterative process may be initiated to come up with a final conclusion through a convergence of opinions and explanations with the involvement of the parties in dispute.

Being just a first step in evaluation, however, they are not capable of providing an exhaustive evaluation on the topic concerned. Evaluation needs to use objective, fact-based input as one important factor for the evaluation which has to be further investigated and triangulated with qualitative techniques.

All of the selected evaluation methodologies tackle more than one general intervention objectives, or, as they are translated to evaluation activities, the evaluation themes.

The methods have been developed and applied with the purpose to answer evaluation questions under the umbrella of evaluation themes. Therefore it might be obvious that they are linked to one of them, however, planning and implementation are rarely identical.

In all cases examined in this research, methodologies addressed more than one evaluation theme. Following their purpose, there is always at least one evaluation theme they directly target (Effectiveness and efficiency in the case of SPOT, and Efficiency at both PCA and LTA). This high concentration of efficiency might seem surprising first, however, bearing in mind that evaluations (particularly in the mid-term phase) focus on implementation, it is rather obvious. The methodologies might indirectly address also effectiveness (SPOT) and accountability (PCA, LTA) related evaluation questions. As a result, these three methodologies provide a mixture that can be used to carry out objective, quantitative-based evaluations that tackle all relevant evaluation themes (intervention objectives), i.e. effectiveness (along with relevance), efficiency and accountability.

To sum it up, the answer for the hypothesis is that the three evaluation methods described in this dissertation are justifiably capable of providing objective, sound basis for evaluation activities and also address the relevant intervention objectives pre-set in MEANS, 2000.

#### **4.4 Thesis 4 – Method simplicity and robustness**

**Hypothesis 4: The selected evaluation methodologies provide simple<sup>19</sup> and robust answers to evaluation questions.**

The evaluation methodologies selected in this dissertation can also provide simple and robust answers to the evaluation questions they are related to. However, in the case of the PCA methodology this hypothesis proved to be only partly valid.

All evaluation methods detailed previously use complex data through complex transformations to provide simple outputs. It is much more beneficial to have simple and sometimes complex (but not complicated) results which would require much resource for the evaluators to carry out, and additional resources on both the evaluators' side and the stakeholders' side to interpret the process, the data requirements and the results. On the

<sup>19</sup> I pursue the notion of simpleness in the sense of „easy to understand” and „easy to communicate” to broad public.

other hand, these methods provide outputs (used as inputs for the evaluation), but not answers to the evaluation questions as already referred to in the previous section. They are essential building blocks of the evaluation, but their role is to create solid foundations or at least working hypotheses for further investigation.

All three methodologies (SPOT, PCA, LTA) assessed have been found user-friendly with potential results that are easy to communicate and work on with. All of them are fairly simple and powerful. They are not intended to and not designed to cover all evaluation aspects, but they all provide the opportunity of a sound evaluation basis directly proportional to their „costs” i.e. data requirements, preparation requirements and interpretation requirements.

The explanation of the result that PCA only partly validates the hypothesis does not lie in simplicity, but robustness. PCA inevitably uses sampling prior to evaluation, and there is much room for selection bias or misinterpretation in case of this method. Therefore, interpretation of the results have to be accompanied with the assumptions and sampling methodologies used which limits simplicity of communication, though in their absence the results might be used as means of political or other agendas.

## **4.5 Thesis 5 – Method flexibility**

**Hypothesis 5: The selected evaluation methodologies have the flexibility to be customised to the context of the evaluation.**

The methodologies assessed have been developed to act as potential bases for international benchmarking. This means that their parameters either originate directly from Community regulation (which is identical for all member states) or the parameters have the flexibility to find a common platform for comparison. This comparability of results makes them beneficial choices when selecting evaluation methodologies. The flexibility of the methods could have been assessed both in terms of planning (in the preparatory phase of the evaluation, exploring how a certain method could be used bearing in mind their limitations) and implementation (retaining the method during the evaluation but changing its use significantly as a result of changing environment, expectations or data availability).

All three methods (SPOT, PCA, LTA) assessed in this dissertation have shown flexibility and can be tailored to the context (in this sense meaning input and output requirements),

though there is a minimal requirement to the application of the method to yield meaningful answers. This covers both data volume and structure (input) and the potential of the application, e.g. number of stages defined in the SPOT methodology (output). All methods have been tested in terms of their customisation potential prior to evaluation and have all proven themselves.

There have also been instances when the flexibility of the method have been tested in the course of the evaluation. For instance, the mid-term evaluation of the NSRF and the SOP IEC in Romania ran in parallel. As the NSRF is on a higher aggregation level, not all OPs could mutually comply with the data structure required for the SPOT method. Therefore, the NSRF SPOT method had to be altered after a first version with regard to making up a common platform that embraces and represents all OPs in a comparable way. The methods that are more likely to use a sample of projects as an input basis (PCA) are obviously more sensitive (or in other ways, flexible) to the composition of the sample. Therefore, though flexibility is a potentially positive characteristic of an evaluation method, an inadequate choice of sample may bring false, misleading or manipulative results, which is a definite risk of flexibility. As a consequence, it is crucial to highlight that the more flexible the method, the more cautiousness and more advanced user is required to use it in a proper and unbiased way.

Each quantitative methodology assessed has the flexibility in terms of both scope (length, depth of assessment), and economics of scale (in line with the resources available for evaluation). Therefore they are capable of acting as the methodological baselines for small-scale evaluations and adequate fundamental methodologies – in more sophisticated form – for more complex evaluations. The application of the selected mid-term evaluation methodologies in diverse countries showed that methods could have been tailored to the needs of the evaluation (e.g. specific evaluation questions), data availability (fewer stages identified in the process), country context (newly accessing country with limited understanding of certain notions), programme (relevance and detail variable) prior to their use, i.e. in the proposal or preparatory phases of the evaluation. In the course of the assessment of these methodologies, I could identify several benefits and limitations of their application. The fact that method parameters have been tailored to existing data on several occasions during the evaluation is a clear indication of their easy adaptability.

## 5 NEW SCIENTIFIC RESULTS

The *regional programme analysis* (assessment of the EU-12 regions, regional programmes and ROPs) yielded the following scientific results:

1. **Cohesion Policy promoted country-level development, however, it could not counteract growing interregional differences in the EU-12 countries.** Future interventions co-funded by Cohesion Policy instruments should be programmed and implemented with a strong intention and view to decrease the regional polarisation within the EU-12 countries. This includes both the planning of the adequate volume of funds capable of creating a critical mass, the concentration of resources and building interventions on regionally differentiating, competitive abilities and further improving the institutional system responsible for the effective and efficient delivery of the funds.
2. **The diverse approaches to ROP measures in the EU-12 have brought very similar allocation results.** Despite the fact that EU-12 ROPs are very diverse (based on relative weight within NSRF, country ROP approach, and sectoral OP vs. regional OP trade-off), the structure and choice of interventions within is fairly similar. This carries the message that bottom-up planning prevails centralised (EC level or member state level) planning structures, as interventions align to existing capacities and future expectations but still remain development need driven.

The *methodology analysis*, carried out on the 2007-13 NSRF and OP mid-term evaluations in EU-12 member states yielded the following scientific results:

3. **Publishing own methodology developments and testing them in international context (SPOT, LTA and PCA).** This dissertation includes the introduction and analysis of selected methods, which have been developed and continuously tested by myself with the contribution of experts from all over the EU-12 region in the last 10 years. The publication and self-assessment and cross-cutting international comparison of these methods is a valuable add-on to the methodology toolset available to programme evaluators and programmers in the new period as well.

4. **Methods used in 2007-13 mid-term programme evaluations can be transferred** to other intervention levels, fields and sectors including non-EU programme evaluation. The tools described in this dissertation are flexible enough to be capable of enhancing programme implementation and evaluation through better monitoring.
5. **Selected methods in 2007-13 mid-term programme evaluations have not been assessed before from a scientific perspective, therefore advantages and limitations have not been collected, categorised and synthesised in view of the experience gained through their multiple application.** My dissertation should facilitate understanding of the use of these methods with both commencers of evaluations and evaluators, which may result in more focused and more efficient evaluation activities. Also, continued use of these methods in the 2014-20 period may provide comparable results across EU programming periods.

## 6 PRACTICAL APPLICATION OF RESULTS

The scientific results of my research are easily and promptly adoptable and applicable for both potential and practicing evaluators and programmers or commencers of programme evaluations.

*Evaluators* may benefit directly from the results of my research through:

- 1. Description and analysis of methodologies:** Methodologies are described in this dissertation so that even those starting their first evaluation activities may understand how to use them. The international examples make methods even more tangible and understandable and demonstrate their objectivity, simplicity, robustness and flexibility and linkage to general intervention objectives. By collecting and sharing these methodologies in an analytic way, through the presentation of their application and their results in multiple countries, evaluators may use them more widely or with more competence.
- 2. Multiple examples of methodology application:** All methodology descriptions are followed by at least two country and context examples. Thus, use and expectable outcomes of the use of a certain method may be assessed. The international examples may provide inspirations on how to scale, alter, or further develop these methods.
- 3. Summary of methodology features, advantages and drawbacks:** By testing and summarising advantages and limitations of methodologies, evaluators may more consciously choose the selected methods. By closing the gap between expectations and reality, evaluators might be more aware of what they can expect as results from the use of these methods.

From another perspective, *programmers, commencers of evaluations and policy officers* may also benefit from my research.

- 4. Better programming:** By understanding better the dynamics and limitations of evaluation activity, programmers may adjust programming exercises so that future evaluation activities may be simpler, requiring less capacity and expense. By sharing a helicopter view on ROP approaches and practices across the EU-12, policy makers and programmers might be inspired to experiment with different strategies in terms of approach and the trade-off of sectoral and regional programmes.

- 5. Better evaluation orders:** Commencers of evaluators might target their Terms of References and evaluation themes and evaluation questions even better by acknowledging what popular and widely used methods are capable of, and what their limitations are. This may help them set their level of expectations and frameworks of evaluations (i.e. deadlines, capacities, data sources available) even better. ToRs may become more attached to the actual methodology guidance being aware of the finding that many of the mid-term evaluation in 2007-13 was not commenced and as a result, was not conducted entirely in line with the EC guidance.
- 6. Raising awareness on international benchmarking and comparison:** It is an open competition in the EU to allocate and spend funds on objectives supported by the EU. This is similar in concept to a zero sum game, therefore, it is important to use the opportunity to learn from each other and develop. International comparisons in terms of programmes, evaluations, methods, practices or mechanisms provide a potential that has not been entirely tapped yet.
- 7. More focus on results than on performance:** Those commissioning evaluations and writing the Terms of References of evaluations might realise, that it is one of the most important purposes of mid-term evaluations to assess progress in terms of indicators and financials as well. Having determined current status, however, it is important to go into details with efficiency, bearing in mind the order. In “summative view”: first, what we have achieved and second, how we have achieved it, while in a “formative view”: first, what to do to proceed and second, how to enhance steps forward.
- 8. More focus on regional development:** Programme designers should bear in mind that the primary objective of Cohesion Policy and its funds and programmes is to mitigate regional inequalities. All means (Cohesion Policy co-funded programmes) should serve this purpose and all corresponding actions should be aligned to this overarching guiding principle.

## **7 SUMMARY**

My research concentrates on the international comparative assessment of the methods that I developed and tested in mid-term programme evaluation contexts in the EU-12, preceded by EU-12 Cohesion Policy development results and the comparative analysis of EU-12 regional programmes in the 2007-13 EU programming period. The subjects of the research are EU co-financed regional development programmes, evaluations and evaluation methods of the EU-12 countries (EU-10 and EU-2 countries), with emphasis on Bulgaria, Malta and Romania.

The second chapter goes into detail with evaluation theory and practice, ranging from the scientific perspective and the professional communities up to how evaluation on the ground is defined, structured, classified and expected to be carried out. As my research focuses on regional programmes, mid-term evaluations and methodologies, this chapter positions and categorises evaluation activities following the widely acknowledged schools and aspects. Mid-term evaluations are applied in the course of programme implementation, providing a unique opportunity to draw conclusions on what has been achieved so far, whether it is in line with the initial expectations, or it is probable that by the end of the programme objectives will be met. Mid-term evaluations also constitute the basis of assisting the creation of the next programming documents by sharing the experience on the implementation of the ongoing programme.

In the third chapter I described my research objective, which is an analysis of Cohesion Policy regional effects in EU-12 and the comparative analysis of ROPs in the EU-12 with focus on assessing selected methods and tools used for the mid-term evaluations applied in EU member states' NSRFs and OPs in the 2007-13 Community budgetary period. By providing a review of the regional programmes and mid-term evaluation methods, I aim to assist the programmers and evaluators of the 2014-20 period by supporting their choice of methodology and tools; and also to share lessons learnt on the advantages and potential pitfalls of their application.

I developed five hypotheses based on a thorough examination of relevant literature. My first hypothesis focuses on Cohesion Policy results of EU-12 development and regional disparities, while the second hypothesis focuses on exploring the EU-12 ROPs in terms of their approach, relative weight and content in NSRFs, also calling upon their mid-term evaluations. The last three hypotheses address the objectivity and intervention objective

(evaluation theme) linkage of the methods, simplicity and robustness and lastly, flexibility of the methods. In order to challenge the five hypotheses I defined two international comparative analyses: a regional programme analysis and a methodology analysis.

The first step of the research was identifying research subjects and extent. For the regional programme analysis (hypothesis 1-2) I used all EU-12 country ROPs and OPs with regional development measures. For the methodology analysis, the selection mostly included objective criteria (country, programme, evaluation). Out of the EU-12 countries having acceded the EU in 2004 and 2007, seven have been explored in detail. As a result, a Bulgarian and two Romanian programmes and methods have been selected, complemented by two Maltese programmes (being similar in many aspects to CEE countries and providing grounds for comparison) for the methodology analysis and testing hypothesis 3-5.

In the methodology analysis, I described three methods (Simple Progress Overview Tool, Post-contractual Assessment and Lead Time Assessment) in detail and have presented two to four international examples per method to clarify their scope, limitation and potential outcomes of their application. I assessed them against my three working hypotheses on methodology analysis to evaluate their objectivity and evaluation objective linkage, simplicity and robustness and flexibility.

In the regional programme analysis, first, I conducted a Cohesion Policy results analysis to assess the contribution of Cohesion Policy to the main overall development indicators of the EU-12 countries, and the changes of their in-country regional level indicators after accession to the EU. Second, I assessed the context, the programming approach, and the relative financial weight of regional programmes within the NSRFs in the EU-12 (51 OPs altogether).

As per the conclusion of my research, I found that all my hypotheses proved to be valid, with the remark that one of them (simplicity and robustness) was only partly justified in the case of one method (PCA, robustness), as a result of potential sampling issues.

The research yielded new scientific results directly attributable to the hypotheses and beyond. Directly, assessments articulated on mid-term evaluation methods have the

potential to provide assistance to evaluators and programmers of the 2014-20 programming period.

I used the opportunity of this research to conduct an international comparison on some of the recently published methods (SPOT and LTA and PCA) which I actively developed over years and tested in countries across Europe. Indirectly, my research shed light on the fact that countries followed different approaches but gained similar patterns in defining the focal points of regional operational programmes and mid-term programme evaluations and that these methods may be used in contexts different from EU funded programmes as well. My analysis also revealed that mid-term evaluations should focus more on achievements and their enhancement than on the improvement of programme implementation considerations. What is most important of all, is that though Cohesion Policy justifiably promoted country development on aggregate country level, however, it could not counteract sufficiently the growing interregional differences in EU-12 countries, which have grown dramatically over the 2004-2011 period.

I am aware that my research is just a minor step amongst other scientific studies in this field but I sincerely hope that evaluation methods could be a scientific area of higher priority with more, and more active contributions in Hungary.

Following up on my research, I find it important to promote the idea of a comparative analysis on the program evaluation methods used in Hungarian programmes so far. The NSRF and OP evaluations of the 2004-06 and 2007-13 periods total up to around 50, providing excellent grounds for research. As a result of the analysis, a methodology collection and a handbook could be published, which could enhance both programmer and evaluator activities to promote more efficient, more effective and more sustainable program planning and implementation.

## 8 ÖSSZEFOGLALÁS

A kutatásom az EU-12-ben (az Európai Unió 2004-ben csatlakozott tíz, és 2007-ben csatlakozott két tagállamában) a 2007-13 időszakban széles körben alkalmazott közbenső programértékelési módszerek bemutatására és nemzetközi összehasonlító elemzésére fókuszál, melyet az EU-12 Kohéziós Politikai beavatkozások regionális különbségekre gyakorolt hatásainak elemzése és regionális (operatív) programjainak összehasonlító elemzése alapoz meg. A kutatás tárgyai az EU-12 országok, ezen belül is kiemelten Bulgária, Málta és Románia EU társfinanszírozású regionális fejlesztési programjai, értékelései és főként értékelési módszerei.

Disszertációm második fejezete részletesen taglalja az értékelés elméletét és gyakorlatát, a tudományos oldaltól a szakmai szervezeteken keresztül az értékelést végzőkön keresztül bemutatva azt, hogy hogyan definiálhatók, strukturálhatók, csoportosíthatók és kivitelezhetők az értékelések. Tekintettel arra, hogy kutatásom a regionális fejlesztési programokra, közbenső programértékelésekre, és az ezekben használt módszerekre koncentrálok, e fejezet keretében az értékelési tevékenységeket az elfogadott iskolák és szempontok szerint, kritikusan elemeztem és vitattam meg. A közbenső értékelések a programok megvalósítása közben készülnek, mely egyedi lehetőséget nyújt arra, hogy felmérjük az elért eredményeket, megvizsgáljuk, hogy ezek összhangban vannak-e az eredeti várakozásokkal, és lehetséges-e, várható-e, hogy a kitűzött célok a program zárásáig teljesüljenek. A közbenső értékelések emellett jó alapot teremthetnek a következő programozási időszak programtervezésének előkészítésére is, a futó program megvalósítása kapcsán nyert tapasztalatok megosztásával és felhasználásával.

A dolgozat harmadik fejezetében bemutatom a kutatási célokat, mely olyan kiválasztott módszerek és eszközök vizsgálata, melyek a 2007-13-as Európai Unió költségvetési időszakban a tagállamok NSRF és OP közbenső programértékelései során kerültek alkalmazásra. Céloom emellett az EU-12 körében a Kohéziós Politikának a tagállamok általános fejlődéséhez és regionális egyenlőtlenségeinek csökkentéséhez való hozzájárulásának elemzése, továbbá az EU-12 regionális fejlesztési programjainak nemzetközi összehasonlító elemzése, többek között néhány kiválasztott közbenső értékelés felhasználásával. A kutatással és annak eredményeivel támogatni kívánom a 2014-20 időszak értékelőit és programtervezőit, megkönnyítve és tudatosabbá téve

módszerválasztásukat az értékelések és az értékelési módszerek gyakorlati alkalmazási tapasztalatainak, előnyeinek és hátrányainak megosztásán keresztül.

Öt hipotézist fogalmaztam meg a kutatáshoz, melyek közül az első hipotézis a Kohéziós Politika EU-12-k fejlődésére és országon belüli regionális különbségeinek mérséklésére kifejtett hatásait vizsgálja. A második hipotézis az EU-12 ROP-okban alkalmazott stratégiai megközelítést, NSRK-kon belül betöltött relatív pénzügyi súlyt, valamint a tartalmi elemeket vizsgálja, többek között néhány kiválasztott ROP közbenső értékelés segítségével. Az utolsó három hipotézis elfogadott és alkalmazott értékelési módszerek objektivitására és beavatkozási célokhoz (értékelési témákhoz) való kapcsolódására, illetve a módszerek egyszerűségére és robusztusságára, rugalmasságára, testre szabhatóságára vonatkozik. Az öt hipotézis tesztelésére két nemzetközi összehasonlító elemzést definiáltam: egy regionális program elemzést, és egy módszertani elemzést.

Kutatásom előkészítéseként meghatározásra került annak pontos tárgya, a vizsgálat köre. A regionális programokra vonatkozó elemzés (és így az első és második hipotézis) az EU-12 tagállamok és régióik valamennyi ROP-jára és regionális fejlesztési intézkedésére (összesen 51 OP) kiterjed. A módszertani elemzés esetében a kiválasztás kritériumai főként objektív szempontokat (ország, program, értékelés) tartalmaztak. Az EU-12 körben hét országot vizsgáltam meg, melynek eredményeképpen egy bolgár és két román program került kiválasztásra, melyeket két máltai program egészít ki (melyek számos jellemzőjükben hasonlóak a közép- és kelet európai programokhoz, így lehetőséget teremtenek a módszerek összevetésére) a módszertani elemzéshez.

A regionális programok elemzése során a tagállami fejlettségi hátterek bemutatását követően a Kohéziós Politika országos és regionális fejlődésre gyakorolt hatását elemeztem, melyet a tagállamok regionális fejlesztési intézkedéseinek strukturális és formális elemzése egészített ki, néhány regionális fejlesztési program közbenső értékeléseinek felhasználásával. A három kiválasztott közbenső értékelési módszert (Simple Progress Overview Tool – SPOT, Post-contractual Assessment – PCA and Lead Time Assessment – LTA) részletesen bemutatva, kettő-négy ország konkrét példáján keresztül szemléltettem használatukat annak érdekében, hogy tisztázzam és elemezzem korlátaikat és alkalmazásuk várható eredményeit. Minden módszert három módszertani hipotézisem alapján vizsgáltam meg, hogy felmérjem mennyire objektívek, mennyire kapcsolódnak a beavatkozási célokhoz (értékelési témákhoz), mennyire egyszerűek és

roboztusak, illetve rugalmasak és testreszabhatók. Kutatásom következtetései alapján a hipotéziseim igazolást nyertek, azzal a megjegyzéssel, hogy egyikük (roboztusság) kapcsán a hipotézis csak részben nyert igazolást az egyik kiválasztott módszer (PCA) esetében, a potenciálisan fellépő mintavételi kihívásoknak köszönhetően.

A kutatás eredményeképpen előállt új tudományos eredmények közvetlenül a hipotézisekhez kapcsolódnak, részben azokon túlmutatnak. A kutatás eredményei közvetlenül felhasználhatók a 2014-20 időszak programtervezése és programértékelései során. Arra is felhasználtam ezt a lehetőséget, hogy azon közelmúltban publikált módszereim közül minél többet (ilyen a SPOT, az LTA, és a PCA is) nemzetközi összevetésben elemezzek, amelyeket az elmúlt években fejlesztettem és tesztelhettem több európai államban a programértékelések során.

Közvetett módon a kutatás rávilágított arra, hogy a vizsgált országok különböző megközelítést követve is számos hasonlósággal rendelkeztek a regionális operatív programok és közbenső értékelések fókuszterületeinek megfogalmazása terén. A kutatás eredményeképpen az is igazolást nyert, hogy a bemutatott programértékelési módszerek az európai uniós támogatási programoktól eltérő környezetben is rugalmasan alkalmazhatók. Az elemzés rávilágított, hogy a közbenső értékelések során javasolt nagyobb hangsúlyt helyezni az elért eredmények számba vételére és értékelésére, mint a program végrehajtási szempontjainak elemzésére. Talán a leglényegesebb kutatási eredmény, hogy az EU-12-ben a Kohéziós Politika bár bizonyíthatóan hozzájárult az egyes országok fejlődéséhez, az országokon belüli regionális különbségek ellensúlyozásához elégtelennek bizonyult.

Tudatában vagyok annak, hogy a kutatásom csak egy kis lépés a tudományos munkák sorában, és őszintén remélem, hogy az értékelés, és az értékelési módszerek egy, a jelenleginél hangsúlyosabb tudományos területté léphetnek elő, több és aktívabb magyarországi közreműködővel. A kutatásom analógiájára és lehetséges folytatásaként fontos kutatási területnek tartom a hazai programértékelési módszerek összehasonlító elemzésének elvégzését. A 2004-06 és 2007-13 időszak NSRK és OP értékeléseinek száma közelíti az 50-et, mely kellően széles alapot képez. A kutatás eredményeként – többek között – előállítható és publikálható lenne egy olyan módszertani gyűjtemény, mely valamennyi tervező és értékelő munkáját elősegítve lehetővé tenné hatékonyabb, eredményesebb és fenntarthatóbb programok tervezését és megvalósítását.

## 9 REFERENCES

### Scientific publications

1. Bachtler, J. – Wren, C. (2006): Introduction. Evaluation of European Union Cohesion Policy: Research Questions and Policy Challenges. – Regional Studies, Vol. 40.
2. Batterbury, S. C. E. (2006): Evaluation Concept and Methods. Principles and Purposes of European Union Cohesion Policy Evaluation. – Regional Studies, Vol. 40.
3. Boldrin, M. (2006): Regional policies after the EU enlargement. 2006 In: Artis, M. et al. (ed.): The Central and Eastern European Countries and the European Union. Cambridge, Cambridge University Press
4. Bradley J. (2005a): CEE National Development Planning. In: Bradley, J. et al. (ed.): Integration, Growth and Cohesion in an Enlarged European Union. New York, Springer, pp. 275-311.
5. Bradley, J. – Petrakos, G. – Traistaru, I. (2005b): Integration, Growth and Cohesion in an Enlarged European Union. New York, Springer, pp. 1-25.
6. Bradley, J. – Zaleski, J. – Zuber, P. (2005c): The Role of Ex-Ante Evaluation in Bradley, J. – Zaleski, J. – Zuber, P.: The Role of Ex-Ante Evaluation in CEE National Development Planning. A Case Study Based on Polish Administrative Experience. In: Bradley, J. et al. (ed.): Integration, Growth and Cohesion in an Enlarged European Union. New York, Springer
7. Burylo, A. (2006): Evaluation. 2007-2013 Programming Period. Második Magyar Értékelési Konferencia, Budapest
8. Canova, F., (2004): Testing for convergence clubs in income per capita: a predictive density approach”, International Economic Review 45, p.49-77.
9. De la Fuente, A. (2003): The effect of Structural Fund spending on the Spanish regions: an assessment of the 1994-99 Objective 1 CSF, FEDEA working paper 2003-11, Madrid.
10. Dobosi, E. (2003): A regionális elemzések módszertani kérdései. Budapest, Ecostat
11. Dvorková, K. (2014): Sigma Versus Beta-convergence in EU28: do they lead to different results? in Mathematical Methods in Finance and Business Administration. Department of European Integration VSB – Technical University of Ostrava
12. Ederveen, S., J. Gorter, R. de Mooij and R. Nahuis (2002): Funds and Games: The Economics of European Cohesion Policy. CPB Special Publication No. 41. <http://www.enepri.org/files/CPBstudy.pdf>
13. Ekins, P. – Medhurst, J. (2006): The European Structural Funds and Sustainable Development. A Methodology and Indicator Framework for Evaluation. – Evaluation, Vol. 12., No. 4., p. 474-495

14. ESRI (1993): The Community Support Framework – Evaluation and Recommendations for the 1994-1997 Framework , Final Report to the Department of Finance, April 1993
15. European Commission (2014): 6<sup>th</sup> Report on Economic, Social and Territorial Cohesion, July 2014
16. Faragó, L. (2004): A regionalizmus hajtóerői Magyarországon. Tér és Társadalom, 3.sz. p.1–23.
17. Ferry, M. (2013): The Achievements of Cohesion Policy: Evidence and Methodological Challenges from an EU10 Perspective. European Policies research Centre, University of Strathclyde
18. Fitz Gerland J. et al (1999): National Investment Priorities Report for the Period 2000-2006, ESRI Policy research Series, No. 33, March 1999
19. Gabriel Adeleye (1984): The purpose of the dokimasia the University of Maiduguri, Nigeria
20. Gaffey, V. – Nicolas, P. (2005): Building Evaluation Capacity. Presentation. Budapest
21. George, Cowan (1999): A Handbook of Techniques for Formative Evaluation. Routledge, New York
22. Guijt (2014). Participatory Approaches, Methodological Briefs: Impact Evaluation 5, UNICEF Office of Research
23. Guijt, Gaventa (1998). Participatory Monitoring and Evaluation: Learning from Change. IDS Policy Briefing. Brighton
24. Hallie Preskill and Darlene Russ-Eft (2005): Building Evaluation Capacity. Sage Publications, Inc.
25. Hirschman, A.O. (1968): Development Projects Observed. Washington, D.C.: Brookings Institution.
26. Hjertholm, P., & White, H. (2000): Foreign aid in historical perspective: Background and trends. In F. Tarp (Ed.), Foreign aid and development (p. 80-102). New York, NY; Routledge
27. Honohan, P., (ed) (1999): EU Structural Funds in Ireland, A Mid-Term Evaluation of the CSF, 1994 – 99, ESRI Policy Research Series Paper No. 31, July 1999
28. Horváth, Gy. – Lóránd B. (2012): Decentralizáció és gazdasági fejlődés – az olasz példa Közgazdasági Szemle, LIX. évf., 2012. december (p.1273–1298)
29. Horváth, Gy. (2001/a): Európai regionális politika. Budapest-Pécs, Dialóg Campus Kiadó.
30. Horváth, Gy. (2001/b): Regionális támogatások az Európai Unióban. Budapest, Osiris Kiadó.
31. Horváth, Gy. (2004/a): Regionális egyenlőtlenségek Európában. Magyar Tudomány, 165. évf., 9. sz., pp. 962-977.
32. Horváth, Gy. (2004/b): Európai regionális politika. Budapest, Dialóg Campus Kiadó
33. Horváth, Gy. (2005/a): Az uniós tag Magyarország új kohéziós politikájának alapkérdései. In: Sorsfordító esztendő. A 42. közigazdász-vándorgyűlés előadásai. Budapest, TAS -11 Kft., pp. 408-416.

34. Horváth Gy. (2014): A decentralizáció hasznáról in Berki M, Halász L (szerk.) A társadalom terei, a tér társadalma: megújulás és fenntarthatóság a város- és térségfejlesztésben: tanulmánykötet Szirmai Viktória születésnapjára. Budapest: Art Webber Studio Kft., 2014. pp. 25-39. (ISBN:978-963-08-9126-4)
35. Horváth, Gy. (2015): Spaces and places in Central and Eastern Europe : historical trends and perspectives of regional development. Routledge
36. Huey-tsyh, C. (2005): Practical Program Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness. SAGE Publications
37. Ian Gilbert (2010): Why Do I Need a Teacher When I've got Google?: The Essential Guide to the Big Issues for Every 21st Century Teacher
38. Legge, J. (1983): The Chinese Classics (Vol. III), pp. 33–34, The Shoo King: Canon of Shun, Taipei, Southern Materials Centre Inc.
39. Jobbágy, Valér (2014): Az uniós források hazai végrehajtásának és végrehajtó rendszerének értékelése, javaslatok az új programozási ciklusra. Doktori értekezés.
40. Johns, Baltussen, Hutubessy (2003): Programme costs in the economic evaluation of health interventions in Cost Effectiveness and Resource Allocation, Geneva
41. Kessler (2003): Motivate and Reward: Performance Appraisal and Incentive Systems for Business Success. Palgrave, London.p.60.
42. Knüppel, Nikitin (2014): Target-Based Evaluation of Disturbances in Production Systems. Advanced Materials Research, Vol. 1018, pp. 589-596.
43. Loránd, Balázs (2009): Konvergencia és fejlesztéspolitika az Európai Unióban és Magyarországon A területi differenciáltságot befolyásoló hazai fejlesztéspolitika értékelése az I. Nemzeti Fejlesztési Terv pályázati forrásainak allokációja alapján. Doktori disszertáció.
44. Louw, J. (1999): Improving practice through evaluation. In D. Donald, A. Dawes & J. Louw (Eds.), Addressing childhood adversity (pp. 60-73). Cape Town: David Philip.
45. Madhabi Chatterji (2013): A History Of Evaluation 6/26/2013, Teachers College, Columbia University
46. Major, Klára (2001): A nemzetközi jövedelemegyenlőtlenség dinamikája, PHD értekezés, Budapesti Közgazdaságtudományi és Államigazgatási Egyetem Közgazdasági Ph.D. program
47. Molle, W. (2006): The Economics of European Integration. Theory, Practice, Policy. Fifth Edition. Aldershot, Ashgate.
48. Molle, W. (2007): European Cohesion Policy. Routledge, London-New York.
49. Nagy, Attila János (2005): The regional assistance system of the EU and the North Great Plain region.
50. Nemes Nagy, J. (1987): A regionális gazdasági fejlődés összehasonlító vizsgálata. Akadémia Kiadó, Bp.

51. Nemes Nagy, J. (2005): Regionális elemzési módszerek. MTA-ELTE.
52. OECD (1986): Methods and Procedures in Aid Evaluation: A Compendium of Donor Practice and Experience
53. Pálné K. I. (2009): A politika új színtere a régió. Századvég, Pécs-Budapest, p.242.
54. Patton, M. (1997): Utilization-focused evaluation, London: Sage Publications, Inc.
55. Patton, Michael Quinn (2015): Qualitative Research & Evaluation Methods: Integrating Theory and Practice, Fourth Edition, SAGE Publications Inc.
56. Pollitt, C. – Summa H. (1996): Performance audit and evaluation: Similar tools, different relationships? New Directions for Evaluation. Wiley Periodicals, Inc. and the American Evaluation Association 1996/71 p. 29–50
57. Potter, C. (2006). Program Evaluation. In M. Terre Blanche, K. Durrheim & D. Painter (Eds.), Research in practice: Applied methods for the social sciences (2nd ed.) (pp. 410-428). Cape Town: UCT Press.
58. Riedel, R. (2005): European cohesion: Facing the problem of growing regional disparities. DRC.
59. Robert Cassen et al (1986): Does Aid work? Clarendon, Oxford
60. Rossi, Freeman, and Wright (1979): Evaluation: A systematic Approach. Sage Publications Inc.
61. Rubin, Babbie (2010): Essential Research Methods for Social Work. Brooks/Cole Cengage Learning
62. Ryoh Sasaki (2006): A Review of the History and the Current Practice of Aid Evaluation, Western Michigan University
63. S. R. Epstein, Maarten Prak (2010): Guilds, Innovation and the European Economy, 1400–1800.
64. Sanders, F. - Worthen (2004): Program evaluation: Alternative approaches and practical guidelines, 3rd edition. Boston: Allyn and Bacon (Pearson).
65. Scriven, M. (1972): Objectivity and Subjectivity in Educational Research In L.G. Thomas (Ed.) Philosophical redirection of educational research (pp.. 99-142)
66. Scriven, M. (1991): Evaluation Thesaurus. Sage Publications.
67. Scriven, M. (1997): Truth and Objectivity Evaluation, Chapter 33 In Eleanor Chelimsky & William R. Shadish: Evaluation for the 21st Century: A Handbook. ISBN: 9781483348896
68. Smith, Mitton, Cornelissen, Gibson, Peacock (2012): "Using evaluation theory in priority setting and resource allocation", Journal of Health Organization and Management, Vol. 26 Iss: 5, pp.655 - 671
69. Solow, R.M. (1956): "A contribution to the theory of economic growth", Quarterly Journal of Economics 106, 327-368.
70. Terstiege, Stefan (2012): Objective versus Subjective Performance Evaluations. University of Bonn, Bonn Graduate School of Economics

71. UN (1951): Measures for the Economic Development of Under-developed Countries, UN DEA, New York
72. W.W. Rouse Ball (1889): A History of the Study of Mathematics at Cambridg. Cambridge University Press
73. Weiss, C. H. (1998): Evaluation: Methods for Studying Programs & Policies 2nd edition. Prentice Hall. ISBN 013 3097 25 0
74. Williamson, J. (1965): Regional inequality and the process of national development. Economic Development and Cultural Change 14, 3-45.
75. Yuan, Zheng (1994): Local Government Schools in Sung China: A Reassessment. In History of Education Quarterly (Volume 34, Number 2; Summer 1994): p.196-201.

## Official documents

1. American Evaluation Association (AEA) (2004): Guiding Principles For Evaluators. July, 2004.
2. ACIS (2010): A Formative Evaluation of Structural Instruments in Romania, Final Report. July 2010. Ministry of Public Finance Authority for Coordination of Structural Instruments (ACIS)
3. ÁSZ (2007): Az uniós támogatások hazai monitoring és ellenőrzési rendszere működésének ellenőrzéséről. 2007. Budapest, Állami Számvevőszék
4. EC (1998): CSF Evaluation Unit. Review of Ongoing Evaluation Function in the CSF for Ireland, 1994 –1999, October 1998
5. EC (1999a): Commission Regulation (EC) No 438/2001 of 2 March 2001 laying down detailed rules for the implementation of Council Regulation (EC) No 1260/1999 as regards the management and control systems for assistance granted under the Structural Funds
6. EC (1999b): Council Regulation (EC) No 1083/2006 of 11 July 2006 laying down general provisions on the European Regional Development Fund, the European Social Fund and the Cohesion Fund and repealing Regulation (EC) No 1260/1999
7. EC (1999c): Council Regulation (EC) No 1264/1999 of 21 June 1999 amending Regulation (EC) No 1164/94 establishing a Cohesion Fund
8. EC (1999d): Council regulation (EC) No. 1260/1999 of 21 June 1999 laying down general provisions on the Structural Funds. 1999. Official Journal of the European Communities. L 161/1
9. EC (1999e): Regulation (EC) No 1080/2006 of the European Parliament and of the Council of 5 July 2006 on the European Regional Development Fund and repealing Regulation (EC) No 1783/1999
10. EC (1999f): Regulation (EC) No 1783/1999 of the European Parliament and of the Council of 12 July 1999 on the European Regional Development Fund
11. EC (1999g): Regulation (EC) No 1784/1999 of the European Parliament and of the Council of 12 July 1999 on the European Social Fund
12. EC (2000a): Working Paper 1: The New Programming Period 2000-2006: Methodological Working Document. Vademecum on the preparation of Plans and programming documents. European Commission, Directorate General XVI
13. EC (2000b): Working Paper 2: The New Programming Period 2000-2006: Methodological Working Document. Implementation of the Performance Reserve. European Commission, Directorate General XVI

14. EC (2000c): Working Paper 3: The New Programming Period 2000-2006: Methodological Working Document. Implementation of the Performance Reserve. European Commission, Directorate General XVI
15. EC (2000d): Working Paper 4: The New Programming Period 2000-2006: Methodological Working Document. Implementation of the Performance Reserve. European Commission, Directorate General XVI
16. EC (2000e): Working Paper 5: The New Programming Period 2000-2006: Methodological Working Document. Implementation of the Performance Reserve. European Commission, Directorate General XVI
17. EC (2000f): Working Paper 7: The New Programming Period 2000-2006: Methodological Working Document. Ex Ante Evaluation and Indicators for INTERREG (Strand A and B). European Commission, Directorate General XVI
18. EC (2000g): Working Paper 8: The New Programming Period 2000-2006: Methodological Working Document. The Mid-term Evaluation of Structural Fund Interventions. European Commission, Directorate General XVI
19. EC (2000h): Working Paper 8a: The New Programming Period 2000-2006: Methodological Working Document. Mid-term Evaluation of Structural Fund Interventions - INTERREG III. European Commission, Directorate General XVI
20. EC (2000i): Working Paper 8b: The New Programming Period 2000-2006: Methodological Working Document. The Mid-term Evaluation of Structural Fund Interventions - URBAN European Commission, Directorate General XVI
21. EC (2000j): Working Paper 9: The New Programming Period 2000-2006: Methodological Working Document. The Update of the Mid-term Evaluation of the Structural Fund Interventions. European Commission, Directorate General XVI
22. EC (2006a): Working Document No. 1: Indicative Guidelines on Evaluation Methods: Ex-Ante Evaluation European Commission. Directorate General Regional Policy
23. EC (2006b): Working Document No. 2: Indicative Guidelines on Evaluation Methods: Monitoring and Evaluation Indicators European Commission. Directorate General Regional Policy
24. EC (2006c): Working Document No. 3: Commission Methodological Paper giving guidelines on the calculation of public or equivalent structural spending for the purpose of additionality. European Commission. Directorate General Regional Policy
25. EC (2006d): Working Document No. 4: Guidance on the Methodology for carrying out Cost-Benefit Analysis. European Commission. Directorate General Regional Policy

26. EC (2006e): Working Document No. 5: Indicative Guidelines on Evaluation Methods: Evaluation during the Programming Period. European Commission. Directorate General Regional Policy
27. EC (2006f): Working Document No. 7: Reporting on core indicators for ERDF and Cohesion Fund. European Commission, Directorate General Regional Policy
28. EC (2006g): Commission Regulation (EC) No 1828/2006 of 8 December 2006 setting out rules for the implementation of Council Regulation (EC) No 1083/2006 laying down general provisions on the European Regional Development Fund, the European Social Fund and the Cohesion Fund and of Regulation (EC) No 1080/2006 of the European Parliament and of the Council on the European Regional Development Fund
29. EC (2006h): Council Regulation (EC) No 1084/2006 of 11 July 2006 establishing a Cohesion Fund and repealing Regulation (EC) No 1164/94
30. EC (2006i): European Commission: Proposal for a Regulation of the European Parliament and of the Council on the Cohesion Fund and repealing Council Regulation (EC) No 1084/2006
31. EC (2006j): Council Decision of 6 October 2006 on Community strategic guidelines on cohesion (2006/702/EC)
32. EC (2006k): Regulation (EC) No 1081/2006 of the European Parliament and of the Council of 5 July 2006 on the European Social Fund and repealing Regulation (EC) No 1784/1999
33. EC (2007): Cohesion Policy 2007-13 National Strategic Reference Frameworks. Office for Official Publications of the European Communities, Luxembourg ISBN 978-92-79-07465-European Communities
34. EC (2014a): Regulation (EU) No 1303/2013 of the European Parliament and of the Council of 17 December 2013 laying down common provisions on the European Regional Development Fund, the European Social Fund, the Cohesion Fund, the European Agricultural Fund for Rural Development and the European Maritime and Fisheries Fund and laying down general provisions on the European Regional Development Fund, the European Social Fund, the Cohesion Fund and the European Maritime and Fisheries Fund and repealing Council Regulation (EC) No 1083/2006
35. EC (2014b): Regulation (EU) No 1301/2013 of the European Parliament and of the Council of 17 December 2013 on the European Regional Development Fund and on specific provisions concerning the Investment for growth and jobs goal and repealing Regulation (EC) No 1080/2006 laying down tasks, scope, aims, thematic and investment priorities and output indicators for regional development.
36. EC (2014c): Communication From The Commission To The European Parliament, The Council, The European Economic And Social Committee And The Committee Of The Regions - Final Simplification Scoreboard for the MFF 2014-2020. COM(2014) 114 final. Brussels, 3.3.2014
37. EC (2014d): Sixth report on economic, social and territorial cohesion. 2014. July, EC DG Regio (adopted by the European Commission on 23rd July 2014)

38. Eurostat (2008): statistical database. <http://ec.europa.eu/eurostat>
39. MEANS (2000): Evaluating socio-economic programmes: Glossary of 300 concepts and technical terms Volume 6. ISBN 92-828-6626-2 European Commission Luxembourg: Office for Official Publications of the European Communities.
40. MRDPW (2011): Mid-term evaluation of the OPRD by Ministry of Regional Development and Public Works in Bulgaria.
41. OECD (2002): Glossary of Key Terms in Evaluation and Results Based Management. 2002. Organisation for Economic Co-operation and Development.
42. OECD (2008): DAC Principles for Evaluation of Development Assistance - Development Assistance Committee, OECD, 1991, 2008.
43. OPRD MA (2011): Mid-term evaluation of Operational Programme Regional Development (OPRD) 2007-13. Ministry of regional development and public works, Bulgaria
44. PPCD (2010a): Mid-term evaluation of OPI Malta. Planning & Priorities Coordination Department (PPCD) Malta
45. PPCD (2010b): Mid-term evaluation of OPII Malta. Planning & Priorities Coordination Department (PPCD) Malta
46. PPCD (2011): Extract from project implementation database cut-off date: 30.09.2010
47. SOP IEC MA (2010): A Formative Evaluation of Sectoral Operational Programme Improved Economic Competitiveness (SOP IEC) in Romania, Final Report. September 2010. Ministry of Public Finance Managing Authority (MA) for SOP IEC
48. Tavistock (2003): Sourcebook. Evaluating Socio Economic Development. Methods & Techniques. The Guide. Tavistock Institute
49. UMIS (2010): Unified Monitoring Information System (database cut-off date: 31.12.2010)
50. UNDP (2006): 'The Evaluation Policy of UNDP', UNDP Executive Board Document DP/2005/28, May 2006. <http://www.undp.org/eo/documents/Evaluation-Policy.pdf>.
51. UNDP (2009): Handbook on Planning, Monitoring and Evaluating For Development Results. United Nations Development Programme, NY.

## 10 PUBLICATIONS



UNIVERSITY OF DEBRECEN  
UNIVERSITY AND NATIONAL LIBRARY



Registry number: DEENK/217/2015.PL  
Subject: Ph.D. List of Publications

Candidate: Balázs Mezősi  
Neptun ID: IR1R7L  
Doctoral School: Kerpely Kálmán Doctoral School (Regional Sciences)

### List of publications related to the dissertation

#### Foreign language international book(s) (2)

1. Várnai, É., Matolcsy, J., **Mezősi, B.**: EU funds in Central and Eastern Europe: Progress report 2007-08. [s.n.], [s.l.], 57 p., 2009.
2. Várnai, É., Matolcsy, J., **Mezősi, B.**: EU funds in Central and Eastern Europe. [s.n.], [s.l.], 57 p., 2008.

#### Hungarian scientific article(s) in Hungarian journal(s) (2)

3. **Mezősi B.**: Sok kicsi sokra megy.  
*KPMG Energ. Évkv. 2015*, 84-85, 2015. ISSN: 2060-6818.
4. **Mezősi B.**, Árva L., Várhelyi T.: A kormányzati politika felelőssége a tőkebevonás terén.  
*Valóság. 48 (2)*, 1-10, 2005. ISSN: 0324-7228.

#### Foreign language scientific article(s) in Hungarian journal(s) (3)

5. **Mezősi, B.**: A practical means of setting the scene for evaluating aggregated absorption risk in EU co-funded programmes related to project failure.  
*Agrártud. Köz. 59*, 63-66, 2014. ISSN: 1587-1282.
6. **Mezősi, B.**: A simple but powerful tool for the assessment of financial progress in EU co-funded grant programmes - Simple Progress Overview Tool (SPOT).  
*Agrártud. Köz. 59*, 67-72, 2014. ISSN: 1587-1282.



Address: 1 Egyetem tér, Debrecen 4032, Hungary Postal address: Pf. 39. Debrecen 4010, Hungary  
Tel.: +36 52 410 443 Fax: +36 52 512 900/63847 E-mail: [publikaciok@lib.unideb.hu](mailto:publikaciok@lib.unideb.hu), Web: [www.lib.unideb.hu](http://www.lib.unideb.hu)



7. **Mezősi, B.:** Application of lead time analysis as a method in EU co-funded mid-term programme evaluations in 2007-13: the case of Bulgaria and Malta.  
*Agrártud. Közl.* 61, 9-14, 2014. ISSN: 1587-1282.

Foreign language conference proceeding(s) (1)

8. **Mezősi, B.,** Matolcsy, J.: A means for assessment of reasons behind progress in EU Structural Funds co-funded programmes: Simple Progress Overview Tool.  
In: The 9th European Evaluation Society International Conference : Abstract Book. European Evaluation Society, [Prague], 170, 2010.

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.

12 October, 2015



## 11 ACKNOWLEDGEMENTS

I am most thankful for the experts and organisations with whom I had the opportunity and pleasure to develop, test, validate and fine-tune mid-term evaluation methodologies over the past years.

My special thank goes to the following experts cooperating in my research:

*Matolcsy, János*, my former superior, mentor and friend at KPMG, an engineer dedicated to use structures, quantification and easy but robust, objective tools for evaluation. The initial idea of the SPOT method was his invention. I had the pleasure to develop, apply, test, validate and fine-tune it in the course of many evaluations across the CEE region.

*Blink, Dirk* my Team Leader and colleague for 2 years in Romanian evaluation projects. His example led me to go down the scientific path as well as a private sector expert path as they well complement each other.

*Grech, Jan*, Director of Public Sector at KPMG Malta, my project manager for more than a year working on the mid-term evaluations „on the island”. He is the idea inventor of the PCA method described in this study that has been further developed and used in other CEE countries.

*Fitzpatrick, Jim*, a Europe-wide known evaluation guru with whom I had the pleasure to work with for almost 10 years, in many countries in the CEE region. A true believer in evaluation expertise, I learnt much on practical evaluation from him.

Without highlighting more names, I am grateful for all who assisted me in my evaluation methodology expert career and contributed to the development of this study.

I am most thankful for *Prof. Dr. Gyula Horváth* and *Dr. Balázs Lóránd* for all their constructive feedback and guidance on the elaboration and presentation of my research.

Above all, I am most grateful for the technical and scientific guidance and supervision of *Prof. Dr. Janos Nagy*. Without his contribution this dissertation would not have been written.

## **12 DECLARATIONS**

### **DECLARATION**

I have prepared this dissertation at the Kerpely Kálmán Doctoral School of the University of Debrecen with the purpose of obtaining Ph.D. degree.

Budapest, 10 November 2015.

\_\_\_\_\_

Candidate signature

### **DECLARATION**

I hereby certify that Balázs Mezősi Ph.D. candidate has carried out his work under my supervision in 2014-2015 in the Kerpely Kálmán Doctoral School. The candidate has made a determinant contribution to the results achieved in this dissertation. The dissertation is the candidate's individual work. I recommend the acceptance of this dissertation.

Debrecen, \_\_\_\_\_

\_\_\_\_\_

Supervisor signature