



Strathprints Institutional Repository

Polverari, Laura (2016) The new ambitions for 2014-2020 European structural and investment funds evaluation : pouring water in a leaking container? European Structural and Investment Funds Journal, 4 (2). ISSN 2196-8268 (In Press) ,

This version is available at <http://strathprints.strath.ac.uk/57608/>

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<http://strathprints.strath.ac.uk/>) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to Strathprints administrator: strathprints@strath.ac.uk

THE NEW AMBITIONS FOR 2014-20 ESIF EVALUATION: POURING WATER IN A LEAKING CONTAINER?

Laura Polverari

European Policies Research Centre, University of Strathclyde, Glasgow, UK

Introduction

One of the recurring criticisms that have been levied to EU Cohesion policy has been its inability to prove its effectiveness and value for money.¹ These criticisms and the parallel growing pressures to reduce the resources assigned to the policy in recent rounds of budget negotiations, have led to the introduction of changes in the regulatory obligations attached to funding. In the 2014-20 regulations, efforts to improve the policy's effectiveness have primarily related to a set of new obligations intended to improve the results-orientation of programme design and implementation, and to a strengthening of the purposefulness of evaluation activities, including through a shift from evaluating implementation to appraising impacts. This article focusses on the latter theme. It starts with a review of the main evaluation obligations foreseen by the Common Provisions Regulations (CPR)² and with a discussion of the change in evaluation focus advocated by DG Regio. It then reviews the preparation of Evaluation Plans in selected EU programmes³ and these plans' coverage, focus, objectives and resources. To conclude, some reflections are provided on the challenges that the new regulatory framework entails for managing authorities and the degree to which the regulatory innovations discussed are considered by programme authorities to be an improvement compared to the past.

¹ See Bachtler J, Begg I, Charles D and Polverari L (2016, forthcoming) *EU Cohesion Policy in Practice: What Does it Achieve?*, Rowman and Littlefield International and Polverari, L & Bachtler, J with Davies S, Kah S, Mendez, C, Michie, R and Vironen, H (2014) *Balance of Competences Cohesion Review: Literature Review on EU Cohesion Policy*, Final Report to the Department for Business, Innovation and Skills, February 2014.

² 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.

³ The author is Senior Research Fellow at the European Policies Research Centre, University of Strathclyde in Glasgow (UK), email laura.polverari@strath.ac.uk. This article is based on research undertaken for the IQ-Net network. For further details and data see Polverari L (2015) 'The Monitoring and Evaluation of 2014-20 EU Cohesion Policy Programmes', IQ-Net Thematic Paper 36(2). The author would like to thank all the IQ-Net researchers for their research in IQ-Net partner countries and all the IQ-Net partners who participated in the research. IQ-Net is sponsored by 18 Managing Authorities of ESIF programmes across Europe. The content and conclusions of this article do not necessarily represent the views of individual members of the IQ-Net network.

New evaluation obligations

The main tasks in relation to evaluation during the programming period are specified in art. 54 and art. 56 of the Common Provisions Regulations (CPR)⁴ and include the following.

- **Drafting Evaluation Plans** (for each programme or for more than one programme), to be submitted for approval to the Programme Monitoring Committee (PMC) within one year from the adoption of the programme (arts. 114.1 and 110.2 CPR). Content and suggested structure of the Evaluation Plans is illustrated in detail in the European Commission's 'Guidance Document on Evaluation Plans'.⁵ The Plans can be reviewed and amended by the PMC during the entire lifecycle of programmes.
- **Implementing the evaluations foreseen in the Evaluation Plans.** These evaluations can be of at least two types:
 - *implementation evaluations*, focused on how a programme is being managed and delivered;
 - *impact evaluations*, assessing the effectiveness, efficiency and impact of programmes. These impact evaluations should be carried out so as to ensure that 'at least once during the programming period, an evaluation shall assess how support from the ESI Funds has contributed to the objectives of each priority' (art. 56.3 CPR).
- **Discussing evaluation results within the Programme Monitoring Committee** and drawing and implementing the relevant recommendations (art. 56.3 CPR).
- **Disseminating evaluation reports** via programme websites, through the website of the European Commission (to whom programme managing authorities have the obligation to send all evaluation reports undertaken) and, potentially, through other means (art. 54.4 and art. 56.3 CPR).

If the tasks themselves are not too dissimilar to the 2007-13 programming period (except for the obligatory nature of the Evaluation Plan), the new emphasis on impact evaluation in the regulations and in the guidance document for the 2014-20 period is a key innovation compared to previous programming periods:

'To date Cohesion Policy evaluations have tended to focus more on implementation issues than capturing the effects of interventions. *For the 2014+ period, the Commission wishes to redress this balance and encourage more evaluations at EU, national and regional level, which explore the impact of Cohesion Policy interventions on the well-being of citizens, be it economic, social or environmental or a combination of the three. This is an essential element of the strengthened result-focus of the policy.*⁶ (emphasis added)

⁴ Ex ante evaluation requirements are not discussed in this article as they fell outside the scope of the research paper on which this article is based.

⁵ European Commission (2015) *Guidance document on Evaluation Plans. Terms of Reference for Impact Evaluations. Guidance on Quality Management of External Evaluation*, February 2015 available online at http://ec.europa.eu/regional_policy/sources/docoffic/2014/working/evaluation_plan_guidance_en.pdf.

⁶ European Commission (2014) European Commission (2014) *The programming period 2014-2020. Guidance document on monitoring and evaluation, European Cohesion Fund, European Regional Development Funds*,

A new approach to evaluation

The European Commission's guidance note on monitoring and evaluation mentions two main (recommended) approaches for impact evaluation:

- theory-based impact evaluation, and
- counterfactual impact evaluation.

The recommendation of theory-based impact evaluation is particularly novel within EU Cohesion Policy, whereas counterfactual evaluations had already been undertaken, particularly to appraise impacts business aids and training measures, in previous programming periods.

Traditionally, the evaluation of the impact of Cohesion Policy programmes has been carried out with a view to establishing the achievements realised and the effectiveness of programmes (i.e. whether goals and targets were met). In other words, evaluation has been traditionally 'goal-oriented', with programme activities 'evaluated on the basis of whether they help[ed] achieve formally stated goals' whose meaningfulness was 'taken for granted'.⁷ Goal-oriented approaches to evaluation', also beyond Cohesion Policy, have generally been characterised by a summative function (in line with the ideological/cultural setting from which they emerged, the New Public Management agenda). Their main aim has been to draw conclusions about the achievements of programmes, i.e. what they delivered. However, this type of evaluation has been seen to have had relatively limited value in terms of learning and lesson-drawing.⁸ In their practical application, goal-driven evaluations have been frequently hampered by a neglect of causality, and even led to goal displacement in some cases, and to what van Thiel and Leeuw (2002) have called a 'performance paradox', where outcomes have become goals in their own rights, and a focus on outcomes leads to 'everything but an improvement in performance'.⁹ Moreover, goal-oriented evaluation 'is based on the assumption, that the goals of the program are clearly stated, or even identifiable'.¹⁰ However, 'The political process is such that it tends to produce legislation, and therefore programs, that have vague or even contradictory goals'.¹¹ This is all the more true for Cohesion policy, given its multi-sectoral and long-term nature and the multi-level and partnership-based process through which strategic objectives are formulated. And this is exactly where theory-based evaluation comes in. In contrast with traditional impact evaluation, theory-based approaches to evaluation have a more marked formative ambition. It is this formative function that, according to the European Commission (DG Regio), makes theory-based impact evaluation particularly useful for Cohesion Policy. With its emphasis on the 'mechanisms' of change, theory-based evaluation 'pinpoints at which link in a larger causal chain repair work is needed',¹² providing room for the learning that the European Commission views as 'an overarching objective of all evaluations'.¹³ A further important contribution of theory-based evaluation is that, by reflecting on the theory of change of the

Concepts and Recommendations, March 2014, available online at http://ec.europa.eu/regional_policy/sources/docoffic/2014/working/wd_2014_en.pdf, p. 8.

⁷ Dahler-Larsen P (2007) 'Evaluation and Public Management', in Ferlie E, Lynn L E and Pollitt C (2007) *The Oxford Handbook of Public Management*, p. 624.

⁸ *Ibidem*, p. 626.

⁹ *Ibidem*, p. 627.

¹⁰ Peters G. (2015) *Advanced Introduction to Public Policy*, Cheltenham and Northampton: Edward Elgar, p. 132.

¹¹ *Ibidem*.

¹² Dahler-Larsen (2007) *Op. Cit.*, p. 629.

¹³ European Commission (2014) *Op. Cit.*, p. 6.

programme, it can help reveal important assumptions about a programme that may have been implicit,¹⁴ and in this way it can support the understanding of reasons for under-performance, where this is the case. To do so, theory-based evaluation relies on a disparate mix of methods, which comprise literature review and documental analysis, analysis of administrative data, interviews, case studies, surveys, stakeholder workshops, with a view to reconstructing the programme's logic of intervention through a process aimed at achieving triangulation.¹⁵ Specific techniques are also used to disentangle the effects of the intervention/programme under observation from wider, exogenous developments (e.g. contribution analysis and general elimination methodology). Even when, as is often the case, theory-based evaluation is not able to fully disentangle the outcome of a programme or intervention from the wider developments that take place beyond and beside the policy intervention, one of its advantages, according to its advocates, is exactly its ability to bring such wider developments to the fore as contributing factors to the change observed.¹⁶

The second approach, counterfactual impact evaluation (CIE), has been utilised for many years - although not widely - within Cohesion Policy. It is an approach that appraises the impact of an intervention by establishing the situation that would have occurred had the intervention not taken place. This is done by comparing a group of subjects exposed to an intervention with a comparison group which was not (e.g. a group of firms that received support from a certain scheme, with a group of firms that did not, or a group of trained individuals with a group of similar individuals who have not been trained) or examining beneficiaries of an intervention before and after exposure.¹⁷ There are various techniques for counterfactual evaluation (mentioned in the Commission's guidance document on monitoring and evaluation in 2014-20)¹⁸ and different methods that can be followed for the selection of control groups. This approach is recommended for some types of impact evaluation, notably 'relatively homogeneous interventions with a high number of beneficiaries'.¹⁹ A key feature of counterfactual evaluation is that it aims to appraise what has changed as a result of the intervention, but not the causality of this, i.e. the how and why questions. The formative potential of this approach is therefore considered to be more limited in scope than in TBE (and qualitative or mixed-methods case study research) and it is also for this reason that the Commission is recommending that the two approaches are used so as to complement each other.

Implementing both types of impact evaluation can be challenging for managing authorities. Evaluation methods need to be tailored to the evaluation questions to be answered. This requires programme managers to be able to perform a balancing act between selecting the most meaningful evaluation questions and adopting the most suitable evaluation approaches which, combined together, can answer those questions. It implies being able to tailor evaluation methods to the data – general statistics, administrative data and monitoring data – that is available to the evaluators. It presupposes an ability to apply alternative evaluation approaches and techniques in a way that is consistent not just with the goals of the specific evaluation but also with the human resources and capacities available. And, lastly, it involves dealing with the processes of managing the evaluations and their follow-up, processes that

¹⁴ Marra M (2015) 'Cooperating for a more egalitarian society: Complexity theory to evaluate gender equity', *Evaluation*, 21.1, pp. 32-46.

¹⁵ European Commission (2014) *Op. Cit.*

¹⁶ Mayne J (2012) Contribution analysis: Coming of age?, *Evaluation*, 18(3), 270-280, p. 276, quoted in Nacrošis V (2014) Theory-based evaluation of capacity-building interventions, *Evaluation*, 20(1), 134-50.

¹⁷ Evalued: The resource for the evaluation of Socio-Economic Development, September 2013, p. 97.

¹⁸ European Commission (2014) *Op. Cit.*, p. 7.

¹⁹ *Ibidem.*

are rarely linear and straightforward, especially when evaluation is intended as an open process which involves, as in theory-based evaluation, stakeholders.

Thinking strategically about evaluation: the Evaluation Plans

The new evaluation rules require managing authorities to take a more strategic and systematic approach to evaluation. A central element to this ambition are the (now compulsory) Evaluation Plans (EPs). These plans had to be submitted to the Programme Monitoring Committees (PMCs) for approval within one year from the adoption of the programmes (arts. 114.1 and 110.2 CPR) and can have different scope (a single OP or more OPs, one or more funds etc.). At the time in which the research was undertaken (Spring 2015), IQ-Net partner programme authorities had been working on the preparation of these Plans and it was thus possible to review their anticipated coverage, objectives and scope.

Coverage and objectives

When the empirical investigation was undertaken, the state of play amongst IQ-Net partners with the elaboration of EPs was mixed, partly reflecting the varied state of OP approval. Most EPs at the time were still being drafted, with the aim of submission to the PMCs between June and December 2015. Where progress was not advanced, this was sometimes due to the necessity to focus more on the strategic drafting of the programmes and their operational launch or to the delayed approval of OPs.

The emerging Evaluation Plans appeared to have a varying coverage, largely reflecting the institutional settings of a country or strategic or operational choices (e.g. to allow exploiting economies of scale or the achievement of whole-country overviews etc.). Evaluation plans cover variously:

- a single programme, whether single-fund (e.g. in Nordrhein-Westfalen, England and in the French regions) or pluri-fund (as in Pomorskie (Poland) and in Slovenia, where the EP will cover the multi-Fund OPs for 2014-20, which involve ERDF, ESF and the Cohesion Fund);
- a single fund across an entire country – as in Austria, where there will be two national EPs, one for ERDF, managed and coordinated by ÖROK (the Austrian Conference on Spatial Planning) and one for the ESF, under the responsibility of the Federal Ministry of Labour, Social Affairs and Consumer Protection;
- more than one programme within a region – e.g. in Wales (UK), where a single EP will cover five ESI Funds programmes (two ERDF OPs and two ESF OPs, as well as the rural development plan);
- a single, national EP for all ESI Funds – as in Portugal, where there will be a single national EP for ERDF, ESF, EAFRD and EMFF co-funded programmes, or a single, national EP for ERDF and ESF, as in Denmark;
- an EP for the whole Partnership Agreement, plus individual EPs for the OPs – as in the Czech Republic and Slovakia.

In many cases IQ-Net managing authorities already had Evaluation Plans in 2000-06 and 2007-13. These experiences were helpful to the MAs in terms of allowing them to draw lessons for the 2014-20 Plans. In Portugal, for example, the 2007-13 period was a very positive experience, giving evaluation a more prominent role, enhancing coordination among institutions responsible for the implementation/coordination of the Funds, and ensuring a more coherent approach to evaluation. This experience was useful in informing the development of an Evaluation Plan for the 2014-20 programming period, particularly in terms of: (i) identifying evaluation needs, taking into account the usefulness of the policy for the main stakeholders; (ii) ensuring that IT systems collect relevant information for evaluation purposes; (iii) coordinating the time to collect additional information for the evaluation process with a more timely availability of statistical data; and, crucially, (iv) including in the Evaluation Plan the strategy for communicating the results of evaluation exercises (a theme, that by and large, did not emerge particularly strongly from the IQ-Net country research in relation to many other countries). Similarly, in Wales the 2007-13 EP was considered to have been helpful as it focused on both process and impact, while making a clear separation between them. The experience of monitoring and evaluating the 2007-13 programmes allowed the MA to draw a number of lessons in order to 'build on the successes and take any necessary action to improve' the 2014-20 evaluation approach.²⁰ The MA will therefore continue a number of practices that in 2007-13 proved useful for the management, implementation and accountability of the funds, such as the sophisticated and comprehensive reporting system, the collection of ESF-participant-level data, the undertaking of ESF Leavers Surveys, the dissemination of 'WEFO Research Summaries', all of which enhanced the feasibility, reliability, accessibility and usability of evaluation, and contributed to generate a comprehensive evidence base useful both for programme management and transparency towards stakeholders. In addition, the MA plans to improve on past practice, for example by extending the coverage of the 'WEFO Research Summaries' to the Rural Development Plan and extending, where possible, the use of counter-factual evaluation, successfully applied in evaluating the impact of ESF interventions on unemployed and economically inactive individuals, to selected ERDF interventions. A crucial endeavour to allow this will be the collection of comprehensive data on the businesses supported by ERDF from projects: despite setting-up a data collection system for all ERDF-supported businesses and the undertaking of ERDF Business Surveys also during 2007-13, persuading firms to submit data proved problematic. This time around, therefore, all project sponsors and scheme managers will be asked to agree to a detailed monitoring and evaluation plan with WEFO's RTE Team.²¹

Generally speaking, the Evaluation Plans (EPs) are interpreted as flexible and 'living' documents, to allow accommodation of evaluation needs as they emerge and re-calibrating work where this is deemed to be necessary. This may mean that the EPs remain generic regarding the kind of evaluation foreseen in each year, leaving the details to annual evaluation plans, decided annually in dedicated PMC meetings (e.g. in Slovenia).

The guidelines on EP drafting provided by the Commission were deemed helpful by various IQ-Net partners; there were however concerns that the Commission's guidance documents and appraisal by the Commission of the EPs might go beyond the regulations. Some areas of concern were also highlighted in relation to specific issues, for instance the need to include evaluation questions in the EPs, which some IQ-Net authorities considered premature, discrepancies in the guidance provided for

²⁰ Welsh Government (2015) WPMC15(58) Monitoring and Evaluation Strategy: European Structural and Investment Funds 2014-2020, available from <http://gov.wales/docs/wefo/publications/150910pmcpapersseptember2015.pdf>.

²¹ *Ibidem*.

ERDF and ESF, and the requirement to appraise impact on macro-indicators for small programmes. In a few IQ-Net countries (e.g. France, Poland and Slovakia), national coordinating authorities provided additional guidance to increase certainty and improve the utility of the forthcoming evaluation work.

The objectives associated with the draft EPs and the evaluations therein contained appear to be pluri-fold: from identifying evaluation (and thus data) needs in a timely manner; to feeding the 2017 and 2019 Strategic Reports; to contributing to learning on 'what works' and on whether programmes require adjustments; to appraising the added value of projects and types of interventions, and their contribution to programme goals; to gauging the effectiveness, efficiency and impact of programmes and/or individual priorities, or improving the evaluation environment and the know-how for the management and delivery of evaluation, as well as ensuring that the evaluations undertaken are relevant and of good quality.

Resources and assignment of responsibilities

At the time the research was carried out, most IQ-Net partners had not yet finalised the budgets dedicated to their evaluation activities. Where resources had been set aside (indicatively) for evaluation, they varied considerably, ranging from c. €600,000 in Flanders, to c. €38 million in France – with various amounts in-between. It is interesting to note that where evaluation responsibilities are devolved to programmes or regions, such as in France, a considerable degree of variation can be expected in the financial weight assigned to evaluation activities which is not correlated to the financial scale of programmes.

Like in the past,²² IQ-Net partners display different preferences in relation to whether evaluation work will predominantly be carried out in-house or outsourced to external evaluators contracted through tendering procedures (for example in England, where the MA is planning to commission a framework contract, along the model similarly followed by EU institutions – European Commission, European Parliament). In some cases, both approaches will be pursued – for instance in Austria, Czech Republic, Denmark and Spain (e.g. taking advantage of the expertise, in this latter case, of the Institute of Fiscal Studies attached to the Ministry of Finance and Public Administration, which is the Managing Authority of all Spanish ERDF programmes). Nevertheless, the use of external evaluators appears to be the most common approach, which is in continuity with past practice. Nevertheless, in some countries, e.g. in Greece and Portugal, the 2007-13 experience has highlighted that the evaluation market is still too confined and needs to be developed either by building capacity in other companies to better address the needs of Cohesion policy evaluation or by developing in-house capacities to carry out independent evaluations. This is something that is intended to be addressed during the 2014-20 programming period.

The coordination of evaluation activities is generally undertaken by national coordinating authorities or, where these differ, by the managing authorities, generally supported by Evaluation Units, Evaluation Steering Groups, Evaluation Committees or 'teams' (e.g. in Wales, the Research, Monitoring and Evaluation, RTE, Team).

²² Polverari L, Mendez C, Gross F and Bachtler J (2007) 'Making Sense of European Cohesion Policy: 2007-13 Ongoing Evaluation and Monitoring Arrangements', *IQ-Net Thematic Paper 21(2)*, European Policies Research Centre, University of Strathclyde, Glasgow, [http://www.eprc.strath.ac.uk/ignet/downloads/IQ-Net_Reports\(Public\)/ThematicPaper21\(2\)Final.pdf](http://www.eprc.strath.ac.uk/ignet/downloads/IQ-Net_Reports(Public)/ThematicPaper21(2)Final.pdf) .

Balance between implementation and impacts

In a majority of IQ-Net partner programmes, implementation evaluation will dominate the agenda during the first half of the programming period, whilst impact evaluation will be predominant in the second half (e.g. Czech Republic, France, Slovakia, Spain and others). Some IQ-Net programmes, on the other hand, will adopt staggered (Nordrhein-Westfalen (Germany)) or flexible approaches (Wales (UK)), or concentrate evaluation activity in the later part of the programming cycle. The focus of evaluation activities will vary considerably across IQ-Net partner programmes, and at the time in which the research was carried out was largely still under development. Mostly, thematic – rather than programme-wide approaches – will be favoured.

In some cases the balance between implementation and impact evaluation will only emerge during implementation. For instance, in Austria and Slovenia the choice will depend on the evaluation questions that will be established and will be defined on a case by case basis or annually. In some countries – notably in the Czech Republic, Slovakia and Spain – the intention is to realise a balanced approach covering both implementation and impact. In others, on the contrary, a shift towards impact evaluation (compared to 2007-13) will be quite pronounced (in Denmark, Finland, France, Pomorskie (Poland), Slovenia and Spain). Implementation evaluation will continue to dominate in Flanders (Belgium), largely due to the small financial scale of the programme. Nordrhein-Westfalen (Germany), on the other hand, will pursue a more individual approach, i.e. a mix of implementation and project-specific evaluations (geared towards added value and results rather than impact on the regional economy). The view is that impact evaluations would be scarcely meaningful, given the financial size of the programme relative to other policies.

Only a few IQ-Net partner programmes at the time of research had a degree of clarity in relation to the approaches that they would adopt to evaluate impacts and most were not yet clear about the relative weight between counterfactual impact evaluation and theory-based impact evaluation. This balance will likely emerge as the evaluation work unfolds, and in some cases the choice of methods will emerge from the tendering procedures (e.g. in Nordrhein-Westfalen (Germany) and Slovenia). Where there is some more clarity on methods – such as in the Czech Republic, Wales (UK) and Denmark – preferences vary: theory-based and case study approaches are expected to dominate in the Czech Republic, CIE in Wales (UK), and a mix of the two in Denmark.

Irrespective of the preferred approach, MAs are gearing up towards evaluation in all cases, for example by addressing data and capacity issues with improvements to monitoring systems and collaboration with data providers and with dedicated training initiatives (e.g. in Finland). There are nevertheless concerns about the ability to carry out impact evaluation, and these relate to a number of factors, namely:

- the measurability of impact on macro-aggregates when the programmes are a only small portion of wider public policy spending;
- the timetable according to which impacts can be measured, which is difficult to reconcile with the short-term interests of politicians;
- the operational and methodological difficulties entailed by CIE (e.g. in Greece);
- the limited or even absent experience of impact evaluation in 2007-13, with regard to both CIE and to theory-based impact evaluation. Even when impact evaluations were attempted in 2007-13,

these were sometimes found to have been of relatively low quality or improvable, due to the limited expertise and a lack of quality data;

- a lack of generalised evaluation culture (in Greece) that inhibits the use of more sophisticated evaluation methods that require technical expertise;
- the resource intensity of impact evaluation and the difficulty, for small administrations (such as in Flanders (Belgium)), to supply the human resources that are necessary for the sourcing of necessary data and to deal with evaluators; and, lastly,
- the weak resilience of evaluation planning against institutional changes and budget cuts.

Conclusions

By and large, the research undertaken among IQ-Net partner authorities shows that the results-orientation focus of the new regulations and, related, the new evaluation requirements and the shift in emphasis from implementation to impact evaluation are viewed positively amongst programme authorities. These changes are considered to represent a move forward compared to the past programming period when financial absorption, partly linked to the economic crisis, dominated the agenda.

The Evaluation Plans are seen by and large as useful by IQ-Net partners too, for example because they allow for a more strategic and longer-term reflection on the goals of evaluation. In many cases, Evaluation Plans, albeit not compulsory, were already in place in past programming periods, providing a source of learning for programme authorities. It is fair to say that the new evaluation approaches build not only on the new regulatory requirements and Commission guidelines, but also on the lessons learnt by programme authorities and their increased familiarisation with and use of evaluation. However, looking to the future, there are concerns about some components of the EPs, as requested by Commission guidelines, such as the need to specify evaluation questions and methodologies at a too early stage. Based on the experience of negotiating the programmes, there are also concerns in some cases that the Commission might take a too extensive approach in the appraisals of the EPs, going above and beyond what is required by the regulations.

Insofar as it is possible to provide an overview at this stage, the approaches taken to evaluation with regard to the coverage, objectives and focus of the EPs vary, as vary the resources allocated to evaluation and the choices regarding who undertakes evaluation work – whether the MAs or other bodies within the administration in charge of the programmes, or external consultants selected via public procurements. Such decisions often reflect the financial resources and skills available, and the expertise required for different types of evaluation. Whatever the choices, a crucial point is that great emphasis is intended to be placed on the quality control of evaluation outputs.

Whilst impact evaluation will gain prominence on the whole compared to the past, implementation evaluation will continue to be carried out and it is still too early to appraise how impact evaluation will be undertaken in practice, and whether the skills and data in place will be adequate to enable the realisation of methodologically sound evaluations. Based on the information available when the research was carried out, CIE appears to be preferred over theory-based impact evaluation, perhaps due to the higher familiarity by programme managers with this approach, given that it was already

utilised in past programming periods.²³ However, a number of concerns have been expressed by IQ-Net partners about the practical applicability of both CIE and theory-based evaluation approaches and, in many cases, IQ-Net partners are still unsure about the extent to which such methods will be used.

Two themes that did not emerge particularly strongly from the field-research are the dissemination and follow-up of evaluation activities, and the involvement of stakeholders in the design and delivery of evaluation. Apart from being required by the CPR and emphasised in the Commission guidance document,²⁴ these are fundamental aspects of the evaluation cycle: evaluations, after all, are a powerful learning and accountability tools if and only if they respond to actual perceived needs, are followed-up and disseminated.

The CPR and the Commission guidance documents have sought to strengthen evaluation in a number of ways both directly – by incentivising a new emphasis on impact evaluation and new approaches towards this – and indirectly – e.g. through improved goal and target-setting, and introducing an ex ante conditionality on statistical systems and results-indicators. However, the research behind this article has shown that after various cycles of Cohesion policy implementation, and despite sizeable investments to build capacities, in a few countries there are still perceived weaknesses about evaluation capacities and lack of an evaluation culture, and that data limitations are still hampering evaluation. Further, questions can be raised about the extent to which programme evaluation is really independent, whether it is carried out by the programme authorities or the Commission, and whether the current system is actually adequate to ensure this.

It could be argued that the new CPR and Commission guidance notes are addressing past weakness by adding new demands to old ones, without having sufficiently addressed the problems behind the partly ineffectual responses to already existing requirements: which is equivalent to pouring more water into a leaking container. It may thus be time for a comprehensive and without-preconception stock-taking exercise to appreciate what is working, what is not working and what is needed to make evaluation a real programme management and accountability tool from the perspective of programme managers, beneficiaries and wider stakeholders. A reflection on the role, potential and limitations of the current framework to evaluating Cohesion policy and on the best way to exploit the Commission's expertise and position towards this. For example,

- (i) Would it be more fruitful to grant programme authorities freedom about how to evaluate impacts based on *their* experience, preferences and specific needs?
- (ii) Should more integration be pursued between the evaluation of Cohesion policy programmes (and of the various ESI Funds) and that of domestic policies, especially where programmes are too small to affect macro-indicators?
- (iii) How could subsidiarity in this specific aspect of programme management be best attained? Should the role of the European Commission and how it interprets it in practice be revisited? For example, would it be more helpful if the Commission stood back from incentivising the use of selected methodological approaches and rather focussed on supporting programme authorities in the generation of those capacities that would enable *them* to select those approaches that best serve their needs?

²³ Mini-case studies of selected counterfactual evaluations undertaken in relation to Cohesion policy interventions in Czech Republic, Denmark, Portugal, Poland and the UK (Wales) can be consulted from the full paper from which this article draws, i.e. Polverari (2015) *Op. Cit.*, pp. 57-71.

²⁴ European Commission (2014) *Op. Cit.*, pp. 16-17.

- (iv) And, lastly, should the current framework of responsibilities attached to evaluating Cohesion policy programmes be rethought, de-coupling evaluation from management both at programme and EU levels? What could an alternative framework entail?