Policy Evaluation in the EU: The Challenges of Linking *Ex Ante* and *Ex Post* Appraisal

*Stijn Smismans**

The EU's new approach to policy evaluation is characterised by a focus on closing the policy cycle (linking ex ante and ex post appraisal) and by applying evaluation to all types of policy intervention, whether expenditure or regulatory policy. This article analyses the main features and challenges of this new approach. It first studies the conceptual and interdisciplinary challenge of such an encompassing approach to evaluation. It then assesses the new approach in the light of four key objectives of ex ante and ex post appraisal; ensuring evidence and learning; accountability, transparency and participation; policy coherence; and reducing the regulatory burden.

I. Introduction

Over the last two decades, policy evaluation has attracted increased attention (and resources) from policy-makers, practitioners and academic scholars within many developed countries and international organisations.[1] This trend is confirmed by the recent adoption (in 2014) of an encompassing Framework for Regulatory Policy Evaluation by the Organisation for Economic Co-operation and Development (OECD).[2] The European Union (EU) too has strengthened its evaluation capacity over the last two decades, and has made evaluation a key feature of its recent Smart Regulation agenda. Yet, policy evaluation in the EU has hardly been studied.[3] As evaluation is increasingly expected to play a more central role in European governance, such marginalisation cannot be justified. This Special Issue of the EJRR aims to address this gap in the literature and to bring the topic of evaluation closer to the mainstream of EU regulatory studies and EU studies in general.

In this introductory article of the Special Issue I will set out the main features of the EU's new approach to policy evaluation,[4] and will analyse its main challenges for both practitioners and scholars. The new approach to policy evaluation is mainly characterised by applying evaluation to all policy areas (extending it in particular from expenditure to regulatory policies), and by strengthening the link between *ex ante* and *ex post* evaluation. By extending the parameters of evaluation beyond expenditure policy and focusing on the entire policy cycle, the first challenge of this new approach is both a conceptual and interdisciplinary one. Section II analyses how different practitioner and academic communities have dealt with *ex ante* and *ex post* evaluation in a rather siloed fashion while using similar concepts in different ways. In the spirit of an encom-

* Stijn Smismans is Professor of EU law at the School of Law and Politics and Director of the Centre for European Law and Governance (Jean Monnet Centre of Excellence) at Cardiff University. The research leading to this article has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 313642–LASI ("Law, science and interests in European policy-making"). I would like to thank Rachel Minto for research assistance, and Emanuela Bozzini, Claudio Radaelli, Steven Højlund and Lut Mergaert for useful comments on an earlier draft of this article.

1 E.g. for an overview of 21 countries and three international organisations, see Jan-Eric Furubo, Ray C. Rist and Rolf. Sandahl (eds.), *International atlas of evaluation* (New Brunswick: NJ Transaction, 2002).

2 For more detail, see Anne Meuwese in this Special Issue.

3 There are rare exceptions, mainly in the field of structural funds policy, e.g. Elliot Stern, "Evaluation policy in the European Union and its institutions" *New Directions for Evaluation* (2009), pp. 67 *et sqq.*; Carlos Mendez and John Bachtler, "Administrative reform and unintended consequences: an assessment of the EU cohesion policy 'audit explosion'", 18 *Journal of European Public Policy* (2011), pp. 746 *et sqq.*; Julian Hoerner and Paul Stephenson, "Theoretical Perspectives on Approaches to Policy Evaluation in the EU: The Case of Cohesion Policy" 90 Public Administration (2012), pp. 699 *et sqq*. Federico Iannacci, Tony Cornford, Antonio Cordella and Francesco Grillo, "Evaluating monitoring systems in the European social fund context: a sociotechnical approach", 33 *Evaluation Review* (2009), pp. 419 *et sqq.*; and Steven Højlund, "Evaluation use in evaluation systems – the case of the European Commission", 20 *Evaluation* (2014), pp. 428 *et sqq*.

4 The broader historical background of EU policy evaluation is set out further by Højlund in this Special Issue.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

passing approach to evaluation, section II will provide conceptual clarification and set out the interdisciplinary challenges. Section III analyses the main features of the EU's new approach to evaluation. Section IV elaborates the main challenges of this new approach in relation to four key objectives of evaluation: ensuring evidence and learning; accountability, transparency and participation; policy coherence; and reducing the regulatory burden. By analysing the current state of affairs of both *ex ante* and *ex post* evaluation in relation to these four key objectives, the gaps and challenges of the encompassing approach are identified. By focusing critically on the objectives of evaluation, the analysis also addresses the politics of evaluation. The literature on evaluation has been criticised for focusing on assessing the effectiveness of evaluation tools and developing the most appropriate methodology, while eschewing analysis of the interests and politics of policy evaluation.[5] Analysing the objectives of evaluation allows raising questions about the appropriate institutional setting, the main actors, interests and ideology at stake in the EU's new approach to policy evaluation.

II. Conceptual and Interdisciplinary Clarification: Evaluation, Appraisal, Audit and Enforcement

1. The Divide in the Literature Between *Ex Ante* Appraisal and *Ex Post* Evaluation

The European Commission defines evaluation as "a critical evidence-based judgement of whether EU action(s) has met the needs it aimed to satisfy and actually achieved its expected effects."[6] All evaluations are supposed to look at effectiveness (do the verified effects correspond to the original objectives?), efficiency (were the costs justified?), relevance (do the original objectives still correspond to the needs of the EU?), coherence (internally and with other interventions with similar objectives), and EU added value (compared to what could be achieved by Member States). In official EU discourse, "evaluation" normally refers to *ex post* (i.e. retrospective) evaluation, which can be interim (i.e. at the mid-term of an initiative), final (at its conclusion), or *ex post* in the strict sense (which can take place several years after the intervention has finished). The Commission talks about *ex ante* evaluation only in relation to expenditure programmes. *Ex ante* evaluation is legally required under the EU's Financial Regulation to assess the cost-effectiveness of all proposed expenditure programmes/actions for the Union budget.

In general, though, Commission documents on evaluation only deal with retrospective evaluation.[7] Moreover, the Commission's well established system of *ex ante* appraisal for regulatory action and main initiatives, namely the system of integrated impact assessments (IIA), is not usually referred to as evaluation. The divide between *ex ante* impact assessments (of mainly regulatory action), and *ex post* evaluation (of mainly expenditure programmes) is also partially reflected within the organisational structure of the Commission, both at the level of the Secretariat General and the Directorates General (DGs). The Secretariat General has separate units for *ex ante* and *ex post* evaluation, and DGs focusing on regulatory intervention tend to copy this pattern although in DGs focusing on expenditure policies, programme units deal often both with *ex ante* and *ex post* evaluation.

This discursive and institutional divide at the EU level reflects a wider gap between policy and academic communities which deal, on the one hand, with *ex ante* regulatory assessment and, on the other hand with evaluation, mainly focused on *ex post* assessment of, in particular, expenditure projects and programmes. Scholars and practitioners dealing with evaluation constitute their own community, distinct from political science, regulation studies, and even public policy analysis.[8] There is also hardly any link between this literature and EU studies, except for some more recent studies (particularly in the field of

5 Melvin M. Mark, Leslie J. Cooksy and William M.K. Trochim, "Evaluation Policy: an introduction and overview", 123 *New Directions in Evaluation* (2009), pp. 3 *et sqq*.. For a similar critique of the literature on *ex ante* appraisal missing out on the "policy and politics" of appraisal, see John Turnpenny, Claudio M. Radaelli, Andrew Jordan and Klaus Jacob, "The policy and politics of policy appraisal: emerging trends and new direction", 16 *Journal of European Public Policy* (2009), pp. 640 *et sqq*.

6 Commission Communication "Strengthening the foundations of Smart Regulation – improving evaluation", COM(2013) 686 final, at p. 7.

7 With the exception of the specific guidelines for *ex ante* evaluation on expenditure programmes. Commission, "Ex Ante Evaluation. A practical guide for preparing proposals for expenditure programmes", DG Budget, 10 December 2001.

8 See also Emanuela Bozzini and Jo Hunt in this Special Issue.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

cohesion policy).[9] The evaluation community has its own journals such as *Evaluation*, *Evaluation Review*, and *New Directions in Evaluation*, and networks such as the European Evaluation Society. The focus of this community is on developing the best methodology for evaluation, particularly for project and programme evaluations. Writing within this community, Mark et al. criticise that evaluation *policy* (i.e. the way in which government, international organisations or private actors provide a set of rules (formal or informal) that shape evaluation practices) has been understudied in the mainstream evaluation literature. [10]

At the same time, a flourishing, although not very large community of scholars has evolved around issues of *ex ante* appraisal. *Ex ante* regulatory appraisal in particular is a key topic of research both in planning studies and in regulatory studies. While the former also constitutes its own research community, with specially dedicated journals, the latter fits more naturally with mainstream political science and policy analysis, given its focus on regulatory intervention, and placing the topic in the context of wider questions about better regulation and the quality of law-making.[11] The latter perspective has also brought the topic closer to the mainstream of EU studies, with journals such as this one paying considerable attention to *ex ante* impact assessments as a key regulatory instrument of the European regulatory state.

As the EU evolves towards an evaluation approach that encompasses both expenditure and regulatory policy, and aims to link *ex ante* and *ex post* evaluation, it is time to bridge the gap between these two scholarly communities.

Conceptually it is important to note that the *ex post* (mainly expenditure) literature has preferred the concept of evaluation, although appraisal and assessment also figure in it. The *ex ante* (mainly regulatory) literature tends instead to use the concepts of appraisal and assessment, and hardly ever uses the term evaluation. There is no strong semantic reason for this choice. According to the Oxford online dictionary, "appraisal" is "an act of assessing something or someone", while "evaluation" is defined as "the making of a judgment about the amount, number, or value of something; assessment". It might be argued that "evaluation" includes more of a (final) judgement call, and therefore seems more naturally *ex post*. However, both *ex ante* appraisal and *ex post* evaluation mechanisms stress that they only provide the evidence for decisions and judgement calls to be made at a more political level afterwards. At the same time, any gathering of information, whether *ex ante* or *ex post*, will include some assessment and judgement on the relevance of information, etc. From this perspective, this article and this Special Issue use the concepts of evaluation and appraisal as synonymous, preferring a broad definition of evaluation and appraisal as including both *ex ante* and *ex post* processes. A similarly broad use of the concept of evaluation is proposed in the OECD's new Framework for Regulatory Policy Evaluation.

Having proposed a broad definition of evaluation as a way to make the *ex ante* and *ex post* research communities meet, it is also worth indicating that, despite the gap between these research communities, there is considerable common ground in these two categories of literature, both regarding current state of play and challenges. Adelle et al. distinguish between four types of research on policy appraisal.[12] While they focus on *ex ante* appraisal, the same categorisation and conclusions can in fact be applied to research on *ex post* evaluation. Type 1 research deals with the design of appraisal/evaluation, focusing on tools and methods; while type 2 aims to assess the performance of *ex ante* and *ex post* appraisal. Type 1 and type 2 constitute by far the bulk of literature on both *ex ante* and *ex post* appraisal, and are partially produced by practitioners and consultancies, often with the aim to advise policy-making on the best way to implement an appraisal. Type 3 research focuses on evidence utilization and whether appraisal leads to policy change via processes of learning. Type 4 takes the investigation further by addressing the real motivations of policy-makers and the interests at stake in policy evaluation. Unlike types 1 and 2 which

9 Highlighted above, *supra* note 3.

10 Melvin M. Mark, Leslie J. Cooksy and William M.K. Trochim, "Evaluation Policy: an introduction and overview", 123 *New Directions in Evaluation* (2009), pp. 3 *et sqq*.

11 Although the topic is only slowly finding its way to the mainstream of public policy. As Turpnenny et al. note, it is telling that no mainstream public policy textbook covers policy appraisal in much detail. John Turnpenny, Claudio M. Radaelli, Andrew Jordan and Klaus Jacob, "The policy and politics of policy appraisal: emerging trends and new direction", 16 *Journal of European Public Policy* (2009), pp. 640 *et sqq.*, at p. 641.

12 Camille Adelle, Andrew Jordan and John Turnpenny, "Proceeding in parallel or drifting apart? A systematic review of policy appraisal research and practices", 30 *Environment and Planning C: Government and Policy* (2012), pp. 401 *et sqq*.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

for a long time have been inspired by the "technical-rational model" (based on the assumption that sound evidence could be gathered neutrally and then presented to political decision-makers), types 3 and 4 are more inspired by a post-positivist understanding of knowledge production and knowledge use in policy-making. Although types 3 and 4 are even less developed for *ex post* evaluation than for *ex ante* evaluation, there is a striking amount of common ground in the *ex ante* and *ex post* research of these types. In relation to the EU, for instance, Dunlop *et al* identify four main usages of IIAs, based on an analysis of implementation practice,[13] while Højlund identifies 10 different usages of *ex post* evaluation in the EU (five during evaluation, five with the final results of evaluation).[14] Although developed separately, these analyses share common inspiration in the literature on knowledge use, which regularly distinguishes the key categories of knowledge use as instrumental (problem solving), strategic (to defend pre-defined positions) and symbolic (to strengthen legitimacy).[15]

As will be analysed further in section III, the new approach to policy evaluation, such as that advocated by the Commission and reflected in broader evaluation trends nationally and internationally, invites further interaction between the *ex ante* and *ex post* research communities as it poses similar challenges for them.

2. Audit and Enforcement

Having proposed a broad concept of evaluation, it is important to delineate the concept from affiliated terms, one that is more closely related to expenditure evaluation and another that relates more to regulatory evaluation, namely audit and enforcement.

Audit has traditionally focused on financial reporting and control over compliance with the rules, thus ensuring effectiveness of management and internal control systems. The last two decades have been witness to a boom in monitoring and auditing systems in the public sector, whether organised internally or externally. This evolution has been labelled as the emergence of an "audit society".[16] As will be further shown below and in Steven Højlund's contribution to this Special Issue, the EU has considerably strengthened its audit and monitoring regime since the 1990s, particularly in the light of legitimacy concerns. Yet, while increased accountability is welcome in any democracy, some have also argued that the audit society is based on a total lack of trust, which not only undermines learning but in the end even democracy, as democracy is a trust-based system in which the electorate entrusts the making of public policy to its elected representatives.[17] That dense compliance and auditing rules stifle learning is also shown in the EU context.[18] However, more recently there has been a shift from compliance audits to performance audits, or at least performance assessments complementing compliance assessments within the auditing process. Performance audit goes beyond the traditional legality and regularity questions of sound financial management and looks into effectiveness and whether objectives have been achieved. However, this means that the distinction between audit and evaluation has become increasingly blurred. According to the European Commission similar knowledge and skills are involved in performance audit and evaluation, and the main difference is the context and purpose of the activities. Audit remains institutionally implemented by institutions set up for financial auditing and has a stronger focus on methodology to assess the economy, efficiency and effectiveness of how the work was done to achieve the objectives, while evaluation also looks into policy impact, considering why something has occurred and the comparison of alternatives.[19] Exploring this further, in this Special Issue Paul Stephenson analyses how the European Court of Auditors (the key EU institution

13 The four main usages are political, instrumental, communicative and perfunctory. Claire Dunlop, Martino Maggetti, Claudio Radaelli and Duncan Russel, "The Many Uses of Regulatory Impact Assessment: A Meta-Analysis of EU and UK Cases", 6 *Regulation and Governance* (2012), pp. 23 *et sqq*.

14 Steven Højlund, "Evaluation use in evaluation systems – the case of the European Commission", 20 *Evaluation* (2014), pp. 428 *et sqq*.

15 Lorna Schrefler, "Reflections on the different roles of expertise in regulatory policy making", in Monika Ambrus, Karin Arts, Ellen Hey, Helena Raulus (eds), *The Role of 'Experts' in International and European Decision-Making ProcessesAdvisors, Decision Makers or Irrelevant Actors?*, (Cambridge: Cambridge University Press, 2014), pp. 63 *et sqq*..

16 Michael Power ,*The Audit Society: Rituals of Verification*, (Oxford: Oxford University Press, 1997).

17 Jan Klabbers, "The virtues of expertise", in Monika Ambrus, Karin Arts, Ellen Hey, Helena Raulus (eds), *The Role of 'Experts' in International and European Decision-Making Processes. Advisors, Decision Makers or Irrelevant Actors*, (Cambridge: Cambridge University Press, 2014), pp. 82 *et sqq*., at p. 87.

18 Steven Højlund in this Special Issue, and Mendez and Bachtler, "Administrative reform", *supra* note 3.

19 Commission, **"**Public Consultation on Commission Guidelines for Evaluation", November 2013, available on the Internet at <http://ec.europa.eu/smart-regulation/evaluation/consultation/index_en.htm>, (last accessed on 20 May 2014), at p. 13.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

to ensure that the books are sound) has more recently engaged in performance auditing, particularly through the production of "special reports", thus carving itself a place in the EU's institutional landscape of policy evaluation.

While performance audit has blurred the line between the audit and evaluation of EU initiatives involving the Union budget, the assessment of performance of regulatory action has focused on the enforcement of regulatory acts at Member State level. Available data are often limited to information on the transposition of EU directives, while information about implementation on the ground remains scarce. Also, the academic literature on compliance in the EU has focused on the transposition of directives, dealing particularly with the identification of explanatory "goodness of fit" variables at national level to explain success or failure of transposition.[20] However, so far, this literature, like the broader Europeanization literature,[21] has not made a link with the debate on evaluation. It has focused on explaining national diversity in relation to European integration, but has not conceptualised the assessment of national implementation as part of an EU evaluation process that can feed back into new European initiatives. This gap between "compliance" and "evaluation" is not only an academic one, but is equally present in the EU's institutional set up. While evaluation has long been the remit of DG Budget, enforcement has always been the responsibility of the more legally oriented Secretariat General of the Commission, linked with the legal service when enforcement action is taken.

In this Special Issue, Melanie Smith breaks the barrier between enforcement and evaluation. The Commission data gathered on enforcement are indeed a good starting point to build up evaluation capacity in relation to regulatory policy. At the same time, Smith argues that the strengthening of evaluation policy can make the Commission's enforcement policy more accountable.

III. Towards an EU Policy Evaluation Culture

1. Key Features of the New Approach to Evaluation

Evaluation is by no means a novelty in EU policy-making. The EU has engaged in project and programme evaluation for several decades now. However, evaluation has been focused on expenditure policies in particular, such as structural funds, research and development, the Common Agricultural Policy (CAP) and development aid. While initial evaluation practices developed ad hoc within different DGs, the increase in the EU budget and EU expenditure during the 1990s, as well as instances of corruption and the legitimacy crisis related to the resignation of the Santer Commission, led to a more systematic approach to evaluation in the European Commission. Legal requirements and control were tightened to ensure financial accountability.[22]

Evaluation became centrally enshrined in budgetary allocations and the seven year financial programming cycle. Evaluation standards were developed by DG Budget in 1999 (and revised in 2004) to guide DGs in their evaluation work, particularly when outsourcing evaluation to external consultants. All DGs were supposed to develop their evaluation capacity, through the establishment of evaluation units in particular. However, as Steven Højlund argues later in this Special Issue, many DGs (particularly those not involved in expenditure policy) considered evaluation a formality they had to comply with, rather than a useful exercise they could learn from.

More recently, the European Commission has aimed towards a reorientation of its evaluation policy. The reorientation of evaluation policy was first set out in the 2007 Commission Communication "Responding to Strategic Needs: Reinforcing the use of evaluation",[23] and was subsequently given more clout by the embedding of evaluation in the broader Bet-

20 Gerda Falkner, Oliver Treib, Miriam Hartlapp and Simone Leiber, *Complying with Europe. EU harmonisation and soft law in the Member States*, (Cambridge: Cambridge University Press, 2005); Thomas König and Brooke Luetgert, "Troubles with Transposition? Explaining Trends in Member-State Notification and Delayed Transposition of EU Directives", 39 *British Journal of Political Science* (2009), pp. 163 *et sqq.*; Ellen Mastenbroek and Michael Kaeding, "Europeanization Beyond the Goodness of Fit: Domestic Politics in the Forefront", *Comparative European Politics* (2006), pp. 331 *et sqq.*; and Esther Versluis "Even Rules, Uneven Practices: Opening the 'Black Box' of EU law in action", 30 *West European Politics* (2007),pp. 50 *et sqq.*.

21 Keith Featherstone and Claudio Radaelli (eds.), *The Politics of Europeanisation* (Oxford: Oxford University Press, 2003); and Paolo Graziano and Maarten Vink, *Europeanization: new research agendas* (Palgrave, 2008)

22 Højlund in this Special Issue.

23 Commission Communication "Responding to Strategic Needs: Reinforcing the use of evaluation", SEC (2007) 2013.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

ter Regulation agenda. The turn from "Better Regulation" to "Smart Regulation" in 2010[24] was precisely characterised by the argument that "better/smart regulation" should be taken into account throughout the *entire* policy-cycle and not just at the start of it (where most better regulation tools had up to then been focused). This means in particular that *ex post* evaluation should gain a more central place in the policy-making process and should be linked to the *ex ante* assessment of new policy intervention. The centrality of evaluation to the policy-making process in general was exemplified by the fact that the Secretariat General of the Commission, instead of DG Budget, became the leading unit responsible for evaluation in 2009. The Commission's 2013 Communication "Strengthening the foundations of Smart Regulation – improving evaluation"[25] reiterates this approach in catchy terms by talking about the "evaluate first" principle, and promoting an "evaluation culture" in the Commission, while more concrete proposals were to be proposed in new evaluation guidelines. The new draft evaluation guidelines were presented for public consultation in November 2013.[26] At the time of writing this article (January 2015) the final new evaluation guidelines have not yet been made publicly available. The start of the new Juncker Commission in November 2014, as well as the parallel redrafting and consultation process on the new impact assessment guidelines[27] may explain the delay. The draft evaluation guidelines confirm the new, more comprehensive approach to evaluation. Drafted by the Secretariat General of the Commission they also read as a more hierarchical document to be followed by all DGs,[28] compared to the existing evaluation guidelines of 2004, which were drafted by DG Budget as a "Practical Guide for the Commission Services", by "presenting practical solutions and good practices".[29]

The key features of the new approach to evaluation resulting from the three Communications (2007 "Reinforcing Evaluation", 2010 "Smart Regulation", 2013 "Improving evaluation") and draft guidelines (2013), are the following:

1) Evaluation has to be applied to all types of EU intervention; expenditure policy as well as regulatory intervention, including soft law measures. While the idea to apply evaluation beyond expenditure policies goes back more than a decade,[30] it is only in the 2007 and 2010 Communications that concrete measures are proposed for a more systematic application of evaluation to regulatory intervention.
2) The "evaluate first" principle locates evaluation firmly within the policy cycle. New EU intervention can only be taken after an assessment of past action has been made. While evaluation of expenditure policy has long been linked to the seven year financial programme cycle, the 2007 Communication sets the Commission on track to fit evaluation of all its action into its strategic planning and programming cycle. Most importantly, *ex post* evaluation should feed back into the EU system of *ex ante* impact assessments, which has been solidly established since 2003.
3) The Commission's emphasis upon the place of evaluation in the policy cycle goes hand in hand with the embedding of evaluation within the Smart Regulation agenda, and the "REFIT programme" in particular. As part of the Smart Regulation agenda, the Commission initiated a Regulatory Fitness and Performance Programme (REFIT) in December 2012[31] in order to review the entire stock of EU legislation; to identify burdens, inconsistencies, gaps or ineffective measures, with the aim to ensure "a simple, clear, stable and predictable regulatory framework for businesses, workers and citizens."[32]

Evaluation has always had a "value for money" character. Control over the implementation of expenditure policies would ensure a certain level of

24 Commission Communication "Smart Regulation in the European Union", COM(2010) 543 final.

25 Commission Communication "Strengthening the foundations of Smart Regulation – improving evaluation", *supra* note 6.

26 Commission, "Public Consultation on Commission Guidelines for Evaluation", *supra* note 19.

27 Commission, "2014 Revision of the European Commission Impact Assessment Guidelines. Public Consultation document", 1 July 2014 , available at <http://ec.europa.eu/smart-regulation/impact/consultation_2014/index_en.htm> (last accessed on 30 September 2014).

28 Although the guidelines do not provide for any systematic screening of DGs on whether they respect these guidelines, in a way that exists through the Impact Assessment Board.

29 Commission, "Evaluating EU Activities. Practical Guide for the Commission Services", DG Budget, July 2004.

30 E.g. Commission Communication "Focus on results: strengthening evaluation of Commission activities", SEC(2000)1051.

31 Commission Communication "EU Regulatory Fitness", COM (2012) 746 final.

32 Commission Communication "Regulatory Fitness and Performance (REFIT): Results and Next Steps", COM(2013) 685 final, at p. 2.

https://doi.org/10.1017/S186729X00004244 Published online by Cambridge University Press

accountability regarding whether citizens had got what they paid for. However, in the Smart Regulation context, the value for money argument develops from *ex post* accountability to making use of *ex post* evaluation to decide on the desirability of future action, particularly in the context of a regulatory framework that aims to be as "smart" and "thin" as possible. The Commission considers evaluation a "key tool" in its Smart Regulation agenda, not only to ensure better (quality) regulation but also to avoid regulatory burden: "Evaluating the effectiveness and efficiency of EU legislation will improve the quality of policy-making and help to identify new opportunities to simplify legislation and reduce administrative burdens."[33] Or as the 2013 Evaluation Communication states it: "There can be a tendency to look forward and focus on new initiatives. But changes are costly and take time to implement – so they need to be justified and greater attention needs to be paid to looking back before moving forward."[34]

4) Finally, by placing evaluation centrally in the policy cycle for all EU action, *ex post* evaluation should ensure above all *policy* learning, and not just (financial) programme or project learning. Evaluation should not only be the remit of a small number of administrators directly involved in a specific programme or project, but should feed back into the political decision-making process. This underpins the Commission's attention to ensuring better communication and transparency so as to increase the number of actors that can be involved in this learning process, whether stakeholders or other institutional actors.

These key features of the new approach to policy evaluation pose important challenges to the EU's evaluation system as developed so far, while inviting a closer link between the *ex ante* and *ex post* research communities.

First, applying evaluation to all types of policy intervention means a particular challenge for regulatory policy. While *ex post* evaluation is particularly developed for expenditure policy, the tools and methodology for *ex post* regulatory policy are much less readily available. Although lessons can be learnt from expenditure policy, tools and methodology cannot simply been transposed to *ex post* regulatory policy assessment. At the same time, the challenge in terms of tools and methodology is not limited to regulatory *ex post* evaluation. For both *ex ante* and *ex post* appraisal the shift from project and programme assessment to broader policy appraisal invites reflection on new methodology, as existing tools may not be sufficient to draw broader policy conclusions.[35].

Secondly, the "evaluate first" principle and focus on the policy cycle implies a major challenge to current practices. For expenditure programmes, the cyclical process is well established, with *ex ante*, midterm, final and *ex post* evaluations constructed around the Commission's Activity Based Management system and budget cycles. However, even for expenditure policy, data from *ex post* evaluation do not always systematically feed back into the current type of *ex ante* financial evaluation. Moreover, for regulatory policy, the idea of a policy cycle is more an abstract notion than a reality. There is neither a systematic cyclical process, nor a broad availability of *ex post* data on regulatory assessment that could feed into new initiatives.

Finally, the increased focus on policy level appraisal, as well as the embedding of evaluation within the Smart Regulation agenda, make evaluation intrinsically more political. This raises questions for the EU about what it really aims to achieve through evaluation and how it can translate this institutionally. For academic research it raises increasingly type 3 and type 4 research questions about the interests at stake, and the political and strategic use of appraisal.

2. Key Tools Developed so far to Realise the New Approach to Evaluation

While the Commission Communications (set out above) define the key features of the new approach to evaluation, there is a long way to go before such an "evaluation culture" is really established. So far the Commission has focused on three key initiatives:

The first key initiative has been to include evaluation within the Commission's strategic planning and to make the results of *ex post* evaluation more readily available. The 2004 Evaluation Guidelines

33 Commission Communication "Smart Regulation in the European Union", *supra* note 25, at p. 4.

34 Commission Communication "Strengthening the foundations of Smart Regulation – improving evaluation", *supra* note 6, at p. 5.

35 Stern, "Evaluation policy in the European Union", *supra* note 3.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

suggest that each DG adopts an annual evaluation plan and a multi-annual evaluation programme.[36] In practice these are integrated in each DGs annual management plan. An overall future evaluation programme is then set out, previously by DG Budget and now by the Secretariat General, to plan Commission evaluations against the strategic priorities of the Commission. The 2010 Smart Regulation Communication promised to publish planned evaluations of legislative action on a specific website to facilitate input. However, while all DGs now provide an annual evaluation plan, there remains significant room for improved longer-term planning, more transparency and greater advance warning and predictability.[37] An overview of planned evaluations is now available on the Commission's evaluation website,[38] as well as in an annex to the Commission Work Programme as part of the overview of REFIT activities. Yet, some DGs have not looked further than a year ahead. Moreover, the overview does not provide links to information on exact dates for tender, and does not provide opportunities for input. Further information is dispersed on DG websites, but the available information is inconsistent. As far as the publication of *ex post* evaluation reports is concerned, there is diversity in the amount of information available on DG websites. While the creation of a central database of evaluation files on the Commission's evaluation website[39] is an important improvement, the database is not exhaustive. [40]

The second key initiative of the new approach relates to attempts to link *ex post* evaluation and the system of *ex ante* integrated impact assessments. Since its inception in 2003, the system of IIA provided two links with evaluation as traditionally defined within the EU. First of all it clarified the relationship between IIAs and existing practice of *ex ante* evaluation required for expenditure initiatives. When an IIA is applied to certain proposals involving budgetary expenditure, the IIA will incorporate those elements specific to *ex ante* expenditure evaluation, particularly on cost-effectiveness issues, in addition to the full assessment of economic, environmental and social impacts an IIA normally provides.[41] Secondly, it provided a link with *ex post* evaluation by requiring impact assessments to set out indicators for future retrospective evaluations of the new initiative. Vice versa, *ex post* evaluation is expected to rely on IIAs and these indicators to identify why and how an intervention was supposed to work in order to then assess implementation.[42] However, the IIA system provided little incentives to include information from previous *ex post* evaluations in impact assessments. Setting out the data sources that IIAs may use, the 2009 IIA Guidelines state that "information *may* include monitoring or evaluation reports from previous or similar programmes" (stress added). However, there is neither a structural obligation nor a procedural incentive to do so. According to the European Impact Assessment Board (IAB) only around one out of six IIAs in 2013 relied on (or used) *ex post* evaluation results (which is, however, already an improvement compared to around one out of ten in 2010).[43] Given the Smart Regulation focus on the policy cycle, the IAB has now committed to screening more systematically whether IIAs make use of data from retrospective evaluations.[44] The new draft IIA guidelines presented in 2014 provide for a further structural linking of *ex ante* and *ex post* assessment. According to these guidelines "embedded in the policy cycle" has become one of the eight core principles of the IIA system. While confirming the previous commitment that each IIA should provide a monitoring and eval-

36 European Commission, "Evaluating EU Activities", at p. 30.

37 Commission Communication "Strengthening the foundations of Smart Regulation – improving evaluation", *supra* note 6, at p. 6.

38 Commission, "Evaluation", 12 November 2014, available on the Internet at <http://ec.europa.eu/smart-regulation/evaluation/index_en.htm> (last accessed on 20 May 2014).

39 Commission, "Search evaluation results", 24 July 2014, available on the Internet at http://ec.europa.eu/smart-regulation/evaluation/search/search.do (last accessed on 21 January 2015).

40 How representative the database is for the entirety of evaluations is difficult to assess. For sure, the database seems to focus on outsourced evaluations, leaving roughly 20% of internal evaluations uncovered.

41 Commission Communication "Impact Assessment", COM(2002)276. Not all IIAs include such financial *ex ante* evaluation as the initiative may not engage the Union budget. Vice versa, *ex ante* financial evaluation continues to exist as a separate process for expenditure actions for which no IIA is required. However, *ex ante* evaluation is now predominantly conducted in the context of impact assessments. See Commission, "Public Consultation on Commission Guidelines for Evaluation", *supra* note 19, at p. 16.

42 Commission, "Public Consultation on Commission Guidelines for Evaluation", *supra* note 19, at p. 7.

43 European Impact Assessment Board, "Annual Report for 2013", available on the Internet at <http://ec.europa.eu/smart-regulation/impact/key_docs/docs/iab_report_2013_en.pdf> (last accessed on 21 January 2015 at p. 7.

44 European Impact Assessment Board, "Annual Report for 2012", available on the Internet at <http://ec.europa.eu/smart-regulation/impact/key_docs/docs/iab_report_2012_en_final.pdf> (last accessed on 21 January 2015), at p. 27.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

uation framework for the future, it now also states that the "impact assessment report should present, and feed into the analysis, the lessons drawn from any relevant retrospective evaluations, fitness checks, implementation experience and infringement activity. When no retrospective evaluations have been carried out, the impact assessment report should clarify why it is still considered opportune to get ahead with a policy initiative."[45] Yet, unlike the current requirement for each IIA report to include a separate section setting out future evaluation and monitoring indicators, the guidelines do not add an additional requirement for each IIA report to include also a separate section on assessing data from previous retrospective evaluations. This gives the Commission more leeway, and makes procedural review by the IAB on this issue a bit more difficult. It suggests that such data would need to be taken into account where relevant throughout the IIA report, although it is more specifically referred to in relation to the first stage of an IIA, namely the definition of the problem, for which the draft guidelines state that "key input to this assessment will be any retrospective evaluations or fitness checks of relevant frameworks already in place."[46]

The third key initiative is the development of tools that allow evaluation beyond the project and programme level. So far the new toolkit for evaluation at *policy* level has focused particularly upon the evaluation of particular policy areas or industrial sectors and is strongly linked with the Smart Regulation agenda and REFIT programme. Traditionally Commission evaluations have been conducted on individual interventions (programmes, legislative acts), which may themselves have involved a range of projects or actions. The Smart Regulation Communication of 2010, instead, proposed the tool of "fitness checks",[47] which aim at a comprehensive evaluation of a policy area (i.e. the evaluation of a group of related interventions that are linked by a common set of objectives). Fitness checks help to give higher political leverage to evaluation, as they extend beyond the tiny network of a single regulatory intervention. Most importantly, though, they are suitable for the purposes of embedding evaluation within the Smart Regulation agenda. A fitness check should assess whether the regulatory framework for a *policy sector* is fit for purpose and provide the basis for policy conclusions on the future of the relevant regulatory framework. It should identify any excessive regulatory burdens, overlaps, gaps, inconsistencies and/or obsolete measures which may have appeared over time, and help to identify the cumulative impact of legislation.[48] The first fitness checks were carried out as a pilot exercise during the period 2010-2013 for policy areas falling under the remit of five DGs (ENV, EMPL, MOVE, SANCO and ENTR). By the end of 2014, the Commission had carried out or launched 47 fitness checks or evaluations specifically aimed at measuring regulatory burden, most of these in the areas of environment (11), enterprise and industry (8) and employment (5).[49] The new Juncker Commission has planned nine new fitness checks for 2015.[50]

In order to evaluate beyond the project or programme level, the Commission has also run a pilot with a new assessment tool, namely the Cumulative Cost Assessment (CCA).[51] Like fitness checks, CCA aims at evaluation at the policy level and at assessing a regulatory framework. However, unlike fitness checks, the focus is not on the regulatory framework in a particular policy area, but on all regulatory interventions that create costs for a particular sector of industry. Focusing only on measuring costs, CCA does not constitute an evaluation on its own, but is said to provide evidence for evaluation, fitness checks and IIAs.[52] Yet, like fitness checks, it is an evaluation tool that is particularly fitting for the purpose of lightening the regulatory burden under the REFIT programme. After the first pilot (in the sector of steel and aluminium), the new Commission has promised to finish three further CCAs in 2015 (in the sectors of forest, chemical, and ceramics and glass industry).

45 Commission, "2014 Revision of the European Commission Impact Assessment Guidelines", *supra* note 27, at p. 29.

46 Ibid., at p. 10.

47 Fitness checks constitute just one of the tools of the wider REFIT programme.

48 Commission, "Public Consultation on Commission Guidelines for Evaluation", *supra* note 19, at p. 16.

49 Commission Communication "Regulatory Fitness and Performance (REFIT): Results and Next Steps", COM(2013) 685 final, at p. 7.

50 Commission Communication "Commission Work Programme 2015. A new start", COM(2014) 910 final, at Annex 3.

51 For a detailed assessment of the new CCA tool, based on analysis of the two pilot exercises, see Lorna Schrefler, Giacomo Luchetta and Felice Simonelli in this Special Issue.

52 Commission Communication "Regulatory Fitness and Performance Programme (REFIT): State of Play and Outlook", COM(2014) 368 final, at p. 15.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

IV. Objectives of *Ex Ante* and *Ex Post* Evaluation in the Light of Closing the Policy Cycle

1. Overview of Objectives

Table 1 provides a comparative overview of the officially claimed objectives of the EU's systems of *ex ante* and *ex post* evaluation.[53] Column two lists the official objectives of IIA as set out in the 2009 IIA Guidelines, while column three provides the four objectives as set out in the 2013 Draft Evaluation Guidelines.[54] The four categories of key objectives set out in column one constitute a common comparative basis for *ex ante* and *ex post* evaluation, based on both the official documents and the academic literature in the two fields.[55] I distinguish four main categories of key objectives of evaluation, namely: 1) ensuring sound evidence and learning; 2) accountability, transparency and participation; 3) policy coherence; and 4) reducing the regulatory burden.

Policy documents often assume that these different objectives of *ex ante* and *ex post* evaluation are entirely complementary. However, such complementarity cannot be taken for granted.

Firstly, even when dealing with *ex ante* and *ex post* evaluation separately, there may be tensions between these objectives. For instance, even though the European Commission tends to present "learning and accountability" as the key objectives of *ex post* evaluation,[56] trying to achieve one objective may work to the detriment of the other. Based on an extensive range of interviews, Steven Højlund (in this Special Issue) shows how the more centrally controlled and formalised evaluation practice that developed in the Commission during the 1990s and 2000s focused on legal and financial accountability, and weakened rather than strengthened the learning capacity that was present prior to that period. Notably, too strong a monitoring to ensure accountability may indeed stifle the learning processes.

Secondly, while the four identified categories of objectives are relevant for both *ex ante* and *ex post* evaluation, they are so in different ways. As such, an automatic fit cannot be assumed when linking *ex ante* and *ex post* evaluation. For instance, an IA system aimed mainly at ensuring policy coherence by coordinating Commission DGs is likely to be less interested in information from *ex post* assessment than an IA system that is particularly steered towards reducing the regulatory burden. Or still, an *ex post* evaluation geared particularly towards reducing the regulatory burden may not deliver the best evidence for an *ex ante* evaluation that aims at providing the widest possible evidence basis for new policy initiatives.

Hence, setting out the objectives in a comparative way allows for the addressing of the challenges of an evaluation culture that aims to close the policy cycle. I will unpack the challenges relating to each key objective below table 1.

2. Sound Evidence and Learning: Which Evidence for whom?

Both *ex ante* and *ex post* evaluation are expected to provide evidence for policy-making. Both the IIA guidelines and the evaluation guidelines state explicitly that they are a decision-support tool, with the aim

53 Claimed objectives do not necessarily correspond with the (strategic) use of evaluation in practice. See Dunlop et al., "The many uses of RIA", *supra* note 13; and Højlund, "Evaluation use", *supra* note 14; as well as Dunlop and Radaelli in this Special Issue.

54 Both texts have been chosen as they provide the most comprehensive official definition of objectives for each category. For *ex post* evaluation I have relied on the 2013 Draft Guidelines, even if not yet in force, because the 2004 guidelines are clearly out of date and no longer in line with the new approach developed since then. For *ex ante* evaluation I have stuck to the 2009 Guidelines, rather than the 2014 Draft IIA guidelines, as they are not yet in force and do not set out the objectives as clearly as the 2009 guidelines.

55 Ann-Katrin Bäcklund, "Impact assessment in the European Commission - a system with multiple objectives", 12 *Environmental Science and Policy* (2009), pp. 1077 *et sqq*, Gerard G. Rowe, "Tools for the control of political and administrative agents: impact assessment and administrative governance in the European Union" in Herwig C.H. Hofmann and Alexander H. Turk (eds.), *EU Administrative Governance* (Cheltenham: Edward Elgar, 2006), pp. 448 *et sqq,*; Michael Scriven, "Beyond Formative and Summative Evaluation" in Milbrey McLauglin and D.C. Philips (eds), *Evaluation and Education: At Quarter Century* (3 edn, Chicago: University of Chicago Press, 1991); Frans-Bauke Van der Meer and Jurian Edelenbos, "Evaluation in multi-actor policy process: Accountability, learning and co-operation", 12 *Evaluation* (2006), pp. 201 *et sqq,*; Susana Borrás and Steven Højlund, "Evaluation and policy learning: the learners' perspective", 54 *European Journal of Political Research* (2015),pp. 99 *et sqq,*.

56 Commission Communication "Strengthening the foundations of smart regulation – improving evaluation", *supra* note 6, at p. 2.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

Table 1

	EX ANTE (2009 IIA Guidelines)	**EX POST (2013 Draft Evaluation Guidelines)**
Evidence, learning	-helps the EU institutions to design better policies and laws. -facilitates better-informed decision making throughout the legislative process.	**A. To provide timely and relevant advice to decision-making and input to political priority-setting:** Evaluation is a decision-support tool. It does not replace, but aids decision-making, both at a strategic (planning) level, and at the level of the design of a new intervention. It aims to raise the quality of the debate, reinforcing the principles of Smart Regulation and administrative simplification. **B. Organisational learning:** The results of an evaluation should be used to improve the quality of an on-going intervention and to prepare future ones: -can take a wider look at an intervention and its environment and identify not just the areas for improvement but also the positive practices and achievements which should be widely shared and if possible duplicated in other areas. -looks at "unintended" and/or "unexpected" effects.
Accountability, transparency and participation	-helps to ensure that the principles of subsidiarity and proportionality are respected, and to explain why the action being proposed is necessary and appropriate. -improves the quality of policy proposals by providing *transparency* on the benefits and costs of different policy alternatives. -takes into account input from a wide range of external stakeholders, in line with the Commission's policy of transparency and openness towards other institutions and civil society.	**C. Transparency and accountability:** EU stakeholders and citizens have a right to ask the Commission to *give an account* of what was done and achieved, not least because tax payers' money is being used to develop and fund the various interventions. This entitles citizens, stakeholders and parliamentarians to hold the administration to account and to see more clearly whether previous promises have materialised and if not what the likely reasons were and what aspects deserve special attention. Transparency can also help to increase trust, as institutions that are transparent and self-critical tend to be more trusted than institutions which do not produce realistic and objective, detailed and full assessments of the performance of their actions. By publishing evaluation findings, the Commission is publicly taking responsibility for its actions, acknowledging how an intervention is performing and inviting further feedback.
Coherence and choice of policy priorities	-ensures early coordination within the Commission. -helps to ensure coherence of Commission policies and consistency with Treaty objectives (such as the respect for Fundamental Rights) and high level objectives (such as the Lisbon or Sustainable Development strategies).	**D. Efficient resource allocation:** Resources are limited and allocation between interventions or even between the separate elements of an intervention should be based on prioritisation of unmet societal/stakeholder needs. The final allocation may be influenced both by the estimated expectations and by any previous experience in running the same or a similar activity.
Reduce the regulatory burden	-improves the quality of policy proposals by providing transparency on the benefits and costs of different policy alternatives and helping to keep EU intervention as simple and effective as possible.	A. .....reinforcing the principles of Smart Regulation and administrative simplification. D. Efficient resource allocation

https://doi.org/10.1017/S186729X00004244 Published online by Cambridge University Press

of aiding but not replacing decision-making. In the academic literature, the use of knowledge in policy-making has often been framed in terms of (policy) learning. Policy learning can be instrumental (how to improve a particular policy), social (leading to more profound paradigmatic changes of policy ideas) and political (when learning affects power and influence).[57] Dunlop and Radaelli propose a broad definition of learning as "the updating of beliefs based on lived or witnessed experiences, analysis or social interaction".[58]

The Evaluation Guidelines talk explicitly about organisational learning, which can take several forms. The Guidelines suggest it can be used to improve ongoing intervention (which is often the case with cyclical expenditure and programme management) as well as future action. While organisational learning on expenditure and programme management is often reflexive within a small circle of policy officers dealing directly with the particular programme, the Guidelines also suggest a process of organisational learning through the exchange of best practices among policy officers in different areas.

In contrast, the IIA Guidelines do not explicitly use the concept of learning, although "better informed decision-making *throughout* the legislative process" suggests some level of reflexivity. For both the *ex ante* and *ex post* guidelines, providing advice to decision-making is not framed in terms of learning; which seems to suggest that organisational learning is conceived to take place at the level of policy officers, while advice to political decision-makers is not labelled as learning.

The key question, though, is not whether the process is explicitly labelled as "learning",[59] but rather which type of evidence both *ex ante* and *ex post* evaluation are expected to deliver, and whether the evidence gathered allows for a reflexive process in which *ex post* evidence feeds into *ex ante* assessment.[60]

Several elements of "misfit" can be identified between the existing systems of *ex ante* and *ex post* evaluation, which makes reflexivity within the policy cycle challenging.

First of all it is worth remembering that, while the new approach to evaluation stresses the link between *ex post* evaluation and the EU's system of IIAs, *ex post* evidence is only one part of the evidence basis for *ex ante* impact assessments. IIAs aim to assess future impacts, in particular economic, social and environmental impacts. Different sources of evidence will be used for that, namely internal and external studies, information gathered via European agencies and advisory committees, as well as consultation with stakeholders. Evidence about what happened with the implementation of previous initiatives is only one part of the equation. At the same time, *ex post* evaluation is not only geared at feeding information into the system of IIAs. Besides other objectives of *ex post* evaluation, such as accountability (see below), evidence may be gathered in function of (financial) programme or project assessment that is not necessarily linked to the IIA system.[61]

Secondly, the "misfit" between the *ex post* evaluation system and IIA system is partially due to the different legal requirements regarding when an IIA and an *ex post* evaluation need to be adopted. IIAs are required for legislative proposals, as well as non-legislative proposals (such as white papers or expenditure programmes) which set out future policies, and implementing measures with likely significant impacts. Completion of an IIA for the latter is at the discretion of the Commission. Legal requirements for *ex post* evaluation are set at different levels. Article 138 TFEU requires the Commission to adopt an annual evaluation report on the Union's finances based on the results achieved. The Financial Regulation sets additional requirements for expenditure policy, in particular for interventions where spending ex-

57 Claudio M. Radaelli and Claire A. Dunlop, "Learning in the European Union: theoretical lenses and meta-theory", 20 *Journal of European Public Policy*, pp. 923 *et sqq*,, at p. 923.

58 Claire A. Dunlop and Claudio M .Radaelli, "Systematising Policy Learning: From Monolith to Dimensions", 61 *Political Studies* (2013), pp. 599 *et sqq*,, at p. 599.

59 The academic literature clearly relies on a broader conceptualization of policy learning than the official documents, and (*ex ante*) evidence providing to policy makers is considered part of it.

60 Focusing on the type of evidence available and the use of such evidence in policy-making may also be a less tricky research strategy than trying to identify these processes as learning processes. Although Dunlop and Radaelli suggest several avenues in policy learning research (based on an analysis of the existing literature), they also seem to indicate that research framed in terms of knowledge utilisation may be the most promising one. Dunlop and Radaelli, "Systematising Policy Learning", *supra* note 58, at p. 615.

61 For the relationship between different types of evidence and the different goals of *ex post* programme evaluation see Marielle Berriet-Solliec, Pierre Labarthe, and Catherine Laurent, "Goals of evaluation and types of evidence", 20 *Evaluation*, (2014), pp. 195 *et sqq*,

https://doi.org/10.1017/S186729X00004244 Published online by Cambridge University Press

ceeds 5 million €.[62] Additional legal requirements for monitoring and reporting are often set out in European secondary legislation, including for non-expenditure policy. However, while expenditure programmes and projects have a clear period of duration around which a programme of *ex ante*, midterm, final and *ex post* evaluation is built, this established pattern is lacking for regulatory interventions. Although some regulatory interventions have sunset clauses, to which a requirement for retrospective evaluation can be attached, most regulatory intervention has no "expiry date". Such interventions may still include monitoring obligations but their scope is highly variable and may refer to the entire regulatory intervention or a specific part of it. There is no inherent link, in either content or timing, between such monitoring obligations and IIA for new regulatory actions.

Thirdly, (partially as a result of the legal requirements) there is a misfit between the type of information gathered through *ex post* assessment and the information needed for *ex ante* assessment. Long focused on expenditure programmes, *ex post* evaluation is particularly geared towards the provision of information on the cost effectiveness of financial interventions. With respect to regulatory intervention, systematic information is mainly limited to compliance reports, particularly the provision of information on the transposition of directives, while assessment of policy implementation is ad hoc and not in depth. Moreover, even when *ex post* evaluation provides in depth assessment of implementation, it will be in relation to the objectives (or part of them) set out in the initial policy measure. This may not be sufficient to ensure broad policy learning to guide new initiatives (with potentially divergent objectives). Moreover, the system of IIAs requires the systematic *ex ante* assessment of economic, environmental and social impacts of new initiatives. The *ex post* evaluation system instead does not require a systematic assessment of initiatives on their economic, environmental and social consequences. The IIA system thus misses out on what could be a very valuable assessment of past practice on the three criteria specifically required for *ex ante* analysis.

Fourthly, as several contributions to this special issue show,[63] one of the key problems of *ex post* evaluation is to clearly identify the original objectives of an initiative and the benchmarks against which evaluation is possible. The systematic use of IIA should help in this regard, since IIAs need to set out the "general", "specific" and "operational" objectives of a new initiative. "General" refers to Treaty-based goals, "specific" to how the new policy would contribute to a Treaty-based goal, and "operational" to specific targets and deliverables to reach that goal.[64] The requirement of the IIA guidelines to set out indicators for *ex post* evaluation is also a welcome help in this regard. However, while most IIAs set out general and specific objectives, many fail to set out operational objectives,[65] making *ex post* measuring particularly difficult. Nearly all IIAs include a section regarding *ex post* monitoring and evaluation, but only about half include specific indicators, although the trend for inclusion is positive.[66] Also, so far there is no available data on whether proposed *ex post* evaluation indicators set out in IIAs definitely make it to the final proposal and whether they are *de facto* used in retrospective appraisal.

Fifthly, an additional complication results from the fact that several evaluations with different scope, assessing different objectives or indicators of the same initiative (or of multiple similar initiatives) may be adopted in parallel or partially in parallel following different life cycles.[67] Although the Commission's initiative to program evaluations more strategically and more visibly is a valuable step in the right direction, it remains a challenge to gather all relevant *ex post* information for new *ex ante* assessment, especially as the new approach aims in particular at broad policy learning (rather than simply project or programme learning).

Finally, there is an imbalance in experience and available evidence regarding *ex ante* and *ex post* eval-

62 Regulation of the European Parliament and the Council on the financial rules applicable to the general budget of the Union and repealing Council of 25 October 2012, Regulation (EC, Euratom) No 1605/2002, at Chapter 7, Article 30.

63 Emanuela Bozzini and Jo Hunt, and Lut Mergaert and Rachel Minto in this Special Issue.

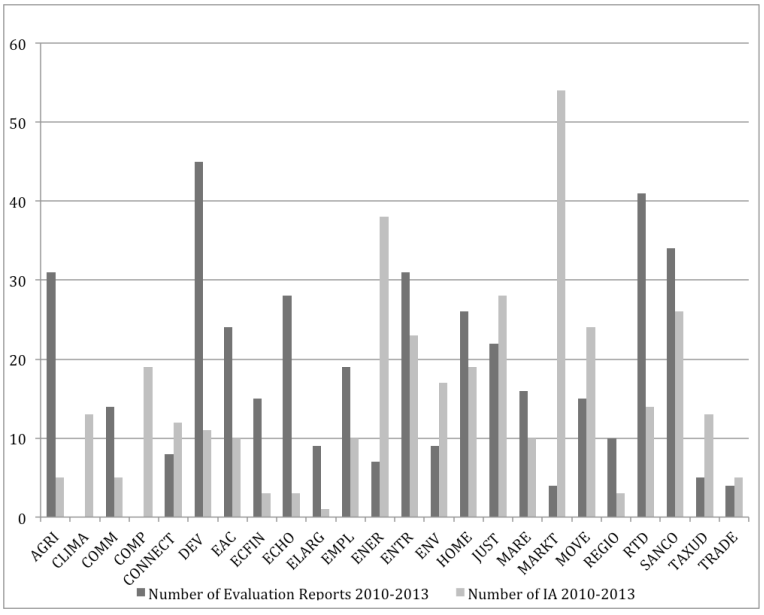
64 Commission, "Impact Assessment Guidelines", SEC(2009) 92.

65 According to a CEPS database of all IIAs adopted between 2003 and 2009, only about 43% of IIAs included operational objectives. The database created by the Centre for European Policy Studies under the supervision of Andrea Renda is not publicly available, but these data were quoted in Giacomo Luchetta, "Impact assessment and the policy cycle in the EU", 3 *European Journal of Risk Regulation* (2012), pp. 561 *et sqq*,, at p. 568.

66 Luchetta, "IA and the policy cycle", *ibid.*, at p. 573.

67 Emanuela Bozzini and Jo Hunt, and Lut Mergaert and Rachel Minto in this Special Issue.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press



Graph 1 - Number of evaluation reports and IAs by DG, 2010-2013

uation in different DGs. Graph 1 compares the total number of IIAs and evaluation reports per DG over a four year period (2010-2013). The data were gathered from the Commission's IIA and evaluation websites, but come with a caveat. While a full list of IIAs is published by the Commission,[68] the public "database on evaluation files"[69] does not systematically cover all evaluation reports. More particularly it focuses on external evaluation reports and tends not to include evaluation reports done entirely within the Commission. Yet, as about 80% of the Commission's evaluation work is outsourced, the available data allow some comparative conclusions between *ex ante* and *ex post* evaluation to be drawn.

Graph 1 shows considerable variation between DGs regarding their use of *ex ante* and *ex post* evaluation. Not surprisingly, DGs whose primary role is in expenditure policies (like DGs AGRI, DEV, REGIO and RTD) adopt a high number of *ex post* evaluations but are modest in their use of IIAs. Vice versa, DGs with primarily regulatory tasks (such as DGs ENER, ENV, and particularly MARKT) have a much stronger tradition of adopting IIAs rather than *ex post* evaluations. This is not to suggest that an encompassing approach to evaluation would necessarily imply that DGs ought to adopt a similar number of *ex post* evaluations and IIAs, and that all DGs should converge to similar practice. However, the graph is indicative of the challenges ahead as it illustrates a misfit between the evidence available through *ex post* evaluation and the inclination for new (regulatory) intervention. With some exceptions (such as DG SANCO, and DG ENTR) DGs which adopt many IIAs and thus propose many new initiatives tend to have relatively limited expertise of and evidence from *ex post* evaluation at their disposal.

It may be concluded that, in terms of evidence and learning, the EU still has some way to go to make *ex ante* and *ex post* evaluation "fit". So far there is little academic research on whether the evidence made available in this process actually leads to learning. A recent study by Boras and Højlund shows how only two types of actors, namely programme units and external evaluators, learned from three expenditure programme evaluations, and that such learning tended to be incremental rather than path-breaking.[70] It remains an open question whether policy learning also occurs in relation to regulatory intervention, and in particular whether it happens at a higher political level in relation to a broader regulatory framework

68 Commission, "2014 impact assessment (IA) reports/IAB opinions , available on the Internet at http://ec.europa.eu/smart-regulation/impact/ia_carried_out/cia_2014_en.htm (last accessed on 13 January 2015).

69 Commission, "Search evaluation results", *supra* note 39.

70 Borrás and Højlund, "Evaluation and policy learning", *supra* note 55.

https://doi.org/10.1017/S186729X00004244 Published online by Cambridge University Press

(in a way as promoted by the new approach to evaluation). The EU's new approach to policy evaluation seems based on a general assumption that *ex post* information will contribute to the sound evidence basis for *ex ante* assessment of new action. However, it has to be acknowledged that different types of evidence lead to learning for different types of actors. Hence, evidence from programme and project evaluation can lead to "administrative" learning at the level of policy officers. Yet, in order to ensure broader policy learning on regulatory intervention, different evidence is required – which to date is scarce – and such evidence has to reach different policy actors. Specifically, it is not enough that information reaches project and programme officers, but evidence provided by research departments dealing with *ex post* evaluation and IIAs must also reach those departments responsible for drafting policy. Moreover, the evidence gathering process may be geared towards objectives other than ensuring the most sound evidence basis for policy-making, such as ensuring accountability, policy coherence or reducing the regulatory burden.

2. Accountability, Transparency and Participation: Who Are and Should be the Actors in Evaluation?

According to Ian Sanderson, the mushrooming of monitoring and evaluation mechanisms in public governance is sign of a process in which governments are turning to evidence of performance for legitimacy since it is no longer guaranteed solely by democratic political processes.[71] Accountability, transparency and participation therefore often appear as objectives of *ex ante* and *ex post* evaluation systems, but in different ways.

Accountability is often identified (together with learning) as the key objective of *ex post* evaluation, both in the literature and in official documents.[72] Accountability has always been core to retrospective evaluation, but it can take different forms.[73] As long as *ex post* evaluation is mainly about auditing, and financial and legal compliance, it is more about "accounting" than "accounting for". What is important is ensuring compliance with the rules by having a threat of sanction rather than political "accounting for" vis-à-vis a broader public or political principal, although it does provide parliament with a means of political control. As evaluation turns increasingly to performance evaluation, it becomes more a means of democratic deliberation allowing "citizens, stakeholders and parliamentarians to hold the administration to account".[74] This explains the Commission's efforts since its 2007 Communication to ensure transparency of evaluation with respect to the other EU institutions (the Parliament in particular) as well as towards citizens. As it states in the 2013 draft guidelines, such transparency can increase trust. However, the focus is on transparency of the outcome of evaluation: by "publishing evaluation findings, the Commission is publicly taking responsibility for its actions, acknowledging how an intervention is performing and inviting further feedback".[75] The invitation for further feedback also suggests a participatory approach to evaluation, yet, only after evaluation reports have been adopted. "The purpose of evaluations, namely to promote accountability/transparency and organisational learning, can only be achieved if the information produced by such evaluations reaches those to whom we are accountable to [*sic*] (general public, parliaments, etc.) or certain intermediaries (journalists) and those who should learn from the results. All evaluation reports of high quality should therefore be disseminated in a manner suited to the different audiences. Active discussion and debate on these findings should be encouraged."[76] The central database of evaluation files available on the Commission's evaluation webpage is a key tool in this regard. To date this database has mainly been limited to providing the evaluation reports. The Commission proposes in the new draft guidelines to publish (in addition to the main evaluation findings) also the evaluation mandate and a Commission response regarding how it will take up these findings. Even if this were realised, these transparency measures function mainly as an accountability tool, while

71 Ian Sanderson, "Evaluation, policy learning, and evidence-based policy-making", 80 *Public Administration* (2002), pp. 1 *et sqq.*, at p,2.

72 Scriven, "Beyond Formative and Summative Evaluation", *supra* note 55; Van der Meer and Edelenbos, "Evaluation in multi-actor policy process", *supra* note 55.

73 See also Steven Højlund in this Special Issue.

74 Commission, "Public Consultation on Commission Guidelines for Evaluation", *supra* note 19, see table 1 above.

75 Commission, "Public Consultation on Commission Guidelines for Evaluation", *supra* note 19, see table 1 above.

76 Commission, "Public Consultation on Commission Guidelines for Evaluation", *supra* note 19.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

participation is conceptualised mainly as an *ex post* process to ensure such accountability rather than as an *ex ante* process ensuring input in evaluation. This fits with a more traditional technical-rational model of evaluation in which evaluation ensures neutral gathering of evidence, which is then presented to political decision-makers to reflect on new action and to the broader public to ensure accountability. However, the new evaluation guidelines also include hints towards a more post-positivist and more participatory approach to the evaluation process itself. Under the title "who contributes information to evaluation?", the draft evaluation guidelines state: "Member States, stakeholders, academics, citizens and a wide range of other parties are involved by providing data and opinion about interventions and wider policies. By *contributing* to, reading and reacting to evaluation reports, they provide further direct input to the decision making system. They play an important part in testing findings and driving independent and impartial evaluations" (stress added).[77] However, the current evaluation system, as well as the new guidelines, remain very modest about how such participation should be organised. The publication of an annual and multi-annual evaluation plan (see above) suggests stakeholders may have a role to play during *ex post* evaluation. However, any concrete mechanisms to ensure such participation are missing.

The balance between accountability, transparency and participation is very different in relation to *ex ante* evaluation. Accountability is less of a core concept of *ex ante* than of *ex post* evaluation, and also the 2009 IIA Guidelines do not use the concept explicitly. However, openness and transparency are often key features of *ex ante* evaluation. Regulatory impact assessments were introduced in several countries during the 1990s in the context of a move towards more "open government".[78] Openness, or transparency,[79] has two dimensions in relation to impact assessments. By way of impact assessments, a regulatory authority has to justify its action, which implies an understanding of openness similar to the idea of accountability. At the same time, ensuring a transparent impact assessment process facilitates participation of stakeholders in *ex ante* evaluation. These two dimensions have been present from the start of the EU's system of IIA. The IIA system was set on track during the early 2000s, in the context of both the Lisbon Strategy and the White Paper on European Governance.[80] The use of IIA as a way to oblige the Commission to provide justification for its initiatives fits the Lisbon Strategy's concern with making the EU more competitive and tackling the regulatory burden. At the same time, the White Paper's concern with more open and participatory European governance provided a second argument to use IIA as a way to make the reasoning and motivations of the Commission more transparent, while at the same time linking IIA to participatory processes. As table 1 above shows, also the 2009 IIA guidelines state clearly that the objective of IIAs is to ensure justification of an intervention (in terms of subsidiarity, proportionality, appropriateness, and the benefits and costs of different policy alternatives), as well as to take into account input from a wide range of external stakeholders. The 2009 guidelines provide further indications about how such participation should be organised, requiring in particular that the 2002 Commission Standards and Principles of Consultation[81] should be respected in this context. The establishment of the system of IIAs has also gone hand in hand with an increased use by the Commission of online consultations, although there is no complete fit between the two processes.[82] Moreover, actual participation patterns in IIAs suggest that the process is more aimed at ensuring policy coherence (see below) than at ensuring the broadest possible participation of stakeholders.[83] Nevertheless, the idea to organise

77 Commission, "Public Consultation on Commission Guidelines for Evaluation", *supra* note 19, at p. 25.

78 Claudio M. Radaelli, "Whither better regulation for the Lisbon agenda?", 14 *Journal of European Public Policy* (2007), pp. 190 *et sqq*,, at p. 192.

79 There are sometimes (subtle) differences between the two concepts as they are used in official documents. Yet, due to limits of space I use them here as synonyms, as is the case in many official documents. For a more nuanced view, see Alberto Alemanno, "Unpacking the principle of openness in EU law: transparency, participation and democracy", 39 *European Law Review* (2014), pp. 72 *et sqq*,; and Stijn Smismans, "Regulating interest group participation in the European Union: Changing Paradigms between transparency and representation", 39 *European Law Review* (2014), pp. 470 *et sqq*,

80 Claudio M. Radaelli, "Whither better regulation", *supra* note 78.

81 Commission Communication "Towards a reinforced culture of consultation and dialogue - General principles and minimum standards for consultation of interested parties by the Commission", COM(2002)704.

82 Still many IIAs do not make use of online consultations, while not all online consultations are used in the context of IIAs. See Emanuela Bozzini and Stijn Smismans, "More inclusive European governance through impact assessments?", *Comparative European Politics*, advance online publication, 9 March, 2015, doi:10.1057/cep. 2015.11.

83 Ibid.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

evaluation as a participatory process is much more enshrined in the *ex ante* than in the *ex post* evaluation system.

This different approach to participation relates to a broader question about who are supposed to be the key actors of *ex ante* and *ex post* evaluation and whether evaluators should be "independent".

Independence is a key argument when *ex post* evaluation is focused on accountability, particularly financial and legal accountability. Within the EU's institutional set-up the Court of Auditors ensures such independence from the Commission in relation to the auditing process. However, the argument about independence extends to most of the EU's *ex post* evaluation system. About 80% of the EU's evaluation work is outsourced to external consultancies. This is due to an issue of resources but also based on the belief that evaluation would be less biased if outsourced to external evaluators than if the Commission assessed its own action, although it has to be acknowledged that the Commission does not entirely lose control over retrospective evaluation if outsourced since it is the Commission which defines the mandate for external evaluation.

The approach is very different in relation to *ex ante* assessment. Although one finds arguments in the literature in favor of external "independent" *ex ante* assessment (particularly as a way to control regulators too keen to act) the EU has not gone this way. It has solidly chosen to organise the IIA system internally within the Commission, although it relies on external sources and documents. Even the quality insurance mechanism of the IIA system, namely the Impact Assessment Board, is organised internally within the Commission. The choice to organise the IIA within the Commission can be understood by the more political nature of *ex ante* evaluation compared to *ex post* evaluation. It is inherently linked with reflecting on different policy objectives and options for new policy intervention. The 2009 IIA Guidelines therefore aim at a balance between gathering "objective" information and providing a picture of the interests at stake. An IIA is expected to provide a balanced overview of available scientific views and expert data, as well as ensuring consultation with "all relevant target groups". The latter comes with the warning that DGs should ensure "peer-reviewing, benchmarking with other studies and sensitivity analysis" to guarantee "the robustness of results" when data are received from stakeholders. The 2014 Draft IIA Guidelines mention "comprehensive", "evidence-based", "unbiased" and "considering a wide and balanced range of stakeholders' views" as fundamental principles of the IIA system. However, organising ex ante evaluation within the Commission includes the risk that IIAs are mainly used by the Commission to justify its preferred policy option.[84] At the same time, organising the evidence gathering function within the bureaucracy facilitates learning.

From this perspective, the objective of an encompassing cyclical approach to evaluation which would engender learning raises several questions for *ex ante* and *ex post* evaluation systems which so far have been based on very different ideas of independence and participation. Firstly, focusing on ensuring independence of *ex post* evaluation may not be the best solution to ensuring that *ex post* results feed back into the policy-making process. Independence is definitely required to ensure financial accountability, but may be less appropriate if the aim is policy learning, in particular for regulatory intervention. At the same time, strengthening the internal evaluation function of the Commission is not in itself a guarantee for such cyclical impact, as one would also need to ensure that evaluation units and IIA units do not operate in isolation from each other. Moreover, even when IIA and *ex post* evaluation are done by the same unit, there is no guarantee that such a research department will have the ability to effectively influence new proposals drafted in the policy department.

Secondly, a more participatory approach to *ex post* evaluation would not only provide valuable information for retrospective evaluation but also create a continuum in the broader set of actors involved both *ex post* and *ex ante*, and thus facilitate the flow of ideas. However, this raises a wider question about who is expected to organise such broader participation. The external consultancies involved in *ex post* evaluation may not have the experience and capacity to organise such participation, and their know-how and methodology may be focused on more narrow assessments, whether financial or cost-benefit assessments. Moreover, wider assessments, both in terms of objectives and actors involved, make the process inherently more political, again raising issues of whether

84 See also Dunlop and Radaelli in this Special Issue.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

these consultancies are the right place to provide such an assessment.

Thirdly, the more political nature of the type of cyclical policy level learning foreseen by the new approach, suggests that a more important role should be envisaged for the European Parliament (EP) in this process. The EP's role in evaluation should not be limited to reading *ex post* reports to sanction the budget. As evaluation is increasingly performance evaluation assumed to feed back into new policy initiatives, the EP should be in a position to judge on the available data and how these feed back into new Commission action. The creation of a Directorate for Impact Assessment and European Added Value in January 2012 has been an important step in the development of such a capacity, although so far its action has focused on the IIA system. In 2014, the Secretary General of the EP therefore also decided to create a unit on *ex post* impact assessment and a unit on policy performance appraisal.[85]

Finally, policy evaluation may be too heavily perceived as a European level game, involving in particular the Commission and an industry of consultancies which operate mainly as European or international businesses. The reality of evaluation in expenditure policy, however, shows a multi-level game, with the EU setting important requirements in terms of reporting and evaluation at the national level. The latter, though, is perceived by Member States as strongly controlling and as not facilitating in terms of learning.[86] The challenge is even bigger in relation to regulatory policy, where there is a clear gap in appropriate infrastructure to gather sufficient and comparable data to feed back from the national to the European level. At the same time, part of the existing EU institutional set-up which can contribute to the evaluation data regarding what happens at a national level, namely (some of) the European agencies, appears only weakly related to the *ex post* and *ex ante* evaluation system.

3. Coherence and Choice of Political Priorities

Both *ex ante* and *ex post* evaluation are assumed to contribute to policy coherence.[87] In line with the predominant expenditure tradition of *ex post* evaluation, the 2013 draft guidelines on evaluation refer to policy coherence in terms of "efficient resource allocation". Coherence is about choosing the most efficient resource allocation based on past experience of the same action or similar action. The comparative basis is limited to the latter rather than referring to the overall objectives of the EU. The aim is internal coherence, that is coherence between each policy and its objective(s),[88] or, most broadly, comparing with other policy initiatives closely related to it. *Ex ante* evaluation instead aims also at external coherence, ensuring consistency of policy action with the multitude of available EU Treaty objectives and other broad EU normative frameworks such as the Europe 2020 Strategy. IIAs have to set the objectives of new intervention in the light of these normative frameworks. Moreover, the systematic screening of economic, social and environmental impacts also steers policy-making in a similar direction, although there is a risk of "steering overload" as the suggested "check list" for assessing impacts is getting ever longer, going from assessing systematically impacts on fundamental rights, to suggestions to assess, among others, territorial impacts, competitiveness, and impacts on micro-enterprises.[89]

Ex ante and ex post evaluation thus have clearly different benchmarks for coherence. The ex ante process is also better institutionally organised to ensure coherence by way of the obligatory creation of an IA Steering Group, bringing together officials from all DGs that may be concerned. Steering Groups for ex post evaluation tend to be more narrowly organised, although they can include external actors, such as external consultants, but generally are less cross-DG than IA Steering Groups.

In the light of a reflexive approach to evaluation, it may be asked whether *ex post* evaluation can deliver more to ensure policy coherence. *Ex post* evaluation is based on assessing outcome in relation to the

85 European Parliament, "European Parliament Work in the fields of ex ante impact assessment and European added value. Activity Report for June 2012-June 2014", European Parliamentary Research Service.

86 Mendez and Bachtler, "Administrative reform", *supra* note 3.

87 The concept of policy coherence has particularly been used by both the OECD and the EU in relation to development policy. It is used here in a more general way to refer to the objective of ensuring coherence within a policy intervention or sector (internal coherence), or ensuring coherence of a policy intervention with other policy objectives of the polity (external coherence).

88 Luchetta, "IA and the policy cycle", *supra* note 65, at p. 564.

89 Commission, "Impact Assessment: Key documents", available on the Internet at <http://ec.europa.eu/smart-regulation/impact/key_docs/key_docs_en.htm> (last accessed on 21 January 2015).

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

objectives set at the origin of an initiative, not in relation to wider EU objectives. However, policy coherence might be improved if *ex ante* evaluation relied on more *ex post* data assessing policy outcomes in relation to broader EU policy frameworks and "governance architectures"[90] such as Europe 2020. This could be done by increasingly setting the objectives of new policy initiatives in light of these broader frameworks, so they could also be assessed on this basis at the level of *ex post* evaluation. Or *ex post* evaluation may engage *de officio* in such broader appreciation, which may suit an evaluation trend that goes from compliance to performance assessment. However, this has two significant inherent risks. *Ex post* evaluation is a difficult and costly task even for policy initiatives with relatively narrow objectives, particularly if the aim is performance evaluation and also the assessment of benefits and not only costs. Indeed, the broader the objectives of an initiative, the bigger the challenge. Moreover, there is a risk that a broad evaluation mandate may be hijacked by assessing technocratically set benchmarks, rather than clearly set Treaty-based objectives formulated by the legislator. Elsewhere we have analysed that there is a tendency in the IIA system to steer policy-making into the direction of rather technocratically framed benchmarks set out, for instance, in the Smart Regulation and Europe 2020 Strategies rather than in function of the EU's constitutional values set out in the Treaties.[91] As the next section will illustrate, the encompassing approach to evaluation is likely to strengthen that process.

4. Reducing the Regulatory Burden

Policy intentions to reduce the regulatory burden are often a key, or even *the* key driver for the introduction of a regulatory impact assessment system. Also at EU level this concern has been at the origin of the creation of the system of IIA, which should be placed in the context of the 2000 Lisbon Strategy which aimed to make Europe the most competitive knowledge based economy in the world. As mentioned above, at the same time it was inspired by concerns about participatory and legitimate governance, expressed in the context of the White Paper on European Governance. The IIA system thus created was more participatory than in most countries, while aimed towards an integrated assessment including assessment of economic, social and environmental impacts. These reflect the original mixed values of the Lisbon Strategy, not only aiming at competitiveness but equally social cohesion and sustainability.

However, the revision of the Lisbon Strategy in 2005 led to an ever stronger focus on competitiveness. As Radaelli has argued,[92] as a consequence, concerns about "quantity" rather than "quality" were brought to the foreground of the Better Regulation agenda and the IIA system. The 2009 IIA guidelines introduced specific sections on "administrative burdens" and on "simplification potential", while a separate "competitiveness proofing toolkit" was introduced as further guidance to the guidelines in 2012.[93]

The initial history of *ex post* evaluation in the Commission has been driven to a lesser extent by concerns about regulatory burden. Although it has always had a strong "value for money" dimension, focus was on *ex post* accountability of expenditure policy. However, as argued above, the new approach to evaluation (extending *ex post* evaluation to regulatory action and focusing on the policy cycle) has turned the "value for money" argument from an *ex post* accountability mechanism into a tool to decide on the desirability of future action. It is the Smart Regulation (2010) agenda and in particular the REFIT programme (2012) (with a strong focus on reducing the regulatory burden) which have propelled evaluation higher up the political agenda. Not surprisingly, the most important new evaluation tools so far (namely fitness checks and CCA) are aimed particularly at addressing the REFIT objectives. The new Juncker Commission has further propelled evaluation to the top of the political agenda by considering it a key tool to address what appears to be its number one political priority: ensuring that the EU is "big on big things

90 Susana Borrás and Claudio M. Radaelli, "The politics of governance architectures: creation, change and effects of the EU Lisbon Strategy", 18 *Journal of European Public Policy* (2011), pp. 463 *et sqq*,

91 Stijn Smismans and Rachel Minto (forthcoming), "Are integrated impact assessments the way forward for mainstreaming in the EU?". See also Dunlop and Radaelli in this special issue, who point to the potentially normatively disturbing finding that IAs develop narratives about values and identities, which are thus developed within bureaucratic documents instead of within constitutional discussions.

92 Radaelli, "Whither better regulation", *supra* note 78.

93 Commission staff working document "Operational guidance for assessing impacts on sectoral competitiveness within the commission impact assessment system. A "Competitiveness Proofing" Toolkit for use in Impact Assessments", SEC(2012) 91 final.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

and small on small things". Promising "a new start" and "making a political priority of lightening the regulatory load",[94] the first action of the new Commission has been to propose a particularly thin work plan, listing only 23 initiatives for 2015, compared to an average of 130 new initiatives each year under the Barroso Commission.[95] In this context, the Commission not only presented a list of 80 acts for withdrawal, but equally a list announcing nine new fitness checks, 44 evaluations of individual regulatory acts, and three new CCAs.[96]

Hence, at the highest political level, evaluation seems to have become primarily a tool to make the EU thinner, based on an assumption that this would be the best way to tackle rising euro-scepticism.

There are several aspects of the current evaluation system which strengthen this push towards using it mainly as an instrument to address the regulatory burden, rather than as a learning system aimed at gathering the best evidence for policy-making in the broadest possible way.

Firstly, unlike impact assessments, which are required to assess the economic, social and environmental impacts of new initiatives (together with impacts on regulatory burden), the general evaluation guidelines do not set specific substantive objectives for *ex post* assessment. *Ex post* assessment always relates to the initial objectives of a policy intervention, which will have to be identified by the evaluation officer on a case by case basis. However, the new evaluation guidelines set out one priority which needs to be assessed more generally, namely "where appropriate, evaluation of regulation should include an assessment of administrative burden, simplification potential, impacts on small and medium-sized enterprises, and the global competitiveness of EU business as part of the analysis of 'efficiency'".[97]

Second, although the Smart Regulation Communication states that individual *ex post* assessments and fitness checks are complementary,[98] the increased attention upon more comprehensive sectoral evaluation includes the risk of turning evaluation increasingly into a function of the agenda to reduce regulatory burden. While individual policy initiatives (whether regulations or programmes) can be clearly assessed on their initial objectives, identifying the initial objectives of a broader regulatory framework is more complex, thus making it more likely for evaluations to focus on externally set evaluation objectives such as impact on administrations, small and medium-sized enterprises, etc.

Thirdly, evaluation is complicated and costly, in particular if an assessment of multiple and broad objectives has to be made. Moreover, measuring benefits is more difficult than measuring costs. Therefore, it may be that the evaluation system will be used particularly in function of one key objective, namely reducing the regulatory burden, with a focus on measuring short-term costs.

V. Conclusion

The EU's intention to create an "evaluation culture" based on a cyclical understanding of policy-making that links *ex post* and *ex ante* evaluation and applies to all types of policy intervention brings with it many challenges. In relation to expenditure policy, the EU has a well-established tradition of cyclical intervention at both programme and project level, including *ex ante* financial evaluation, mid-term and evaluation, followed by a new budget cycle. However, even in these cases, there remain multiple challenges, related to timing (such as *ex post* evaluation coming too late to inspire the next programme round), the difficulty of identifying the initial objectives of initiatives in order to assess them, or the changing nature of these objectives over time. The problems are exacerbated when evaluation is aimed towards performance rather than compliance, particularly when there is a shift from project and programme level evaluation to a broader political evaluation of the regulatory framework. There are many misfits between the key objectives of *ex post* and *ex ante* evaluation and the way they have been institutionalised, going from the gap between the type of evidence gathered *ex post* and the one needed for *ex ante* assessment; the lack of established time frames and the absence of a cyclical process in relation to regulatory intervention; to a different focus on account-

94 Commission Communication "Commission Work Programme 2015. A new start", COM(2014) 910 final, at p. 3.

95 European Voice, "The Companion to the European Commission", February 2015, at p. 24.

96 Commission Communication "Commission Work Programme 2015", *supra* note 94, at Annex 3.

97 Commission, "Public Consultation on Commission Guidelines for Evaluation", *supra* note 19, at p. 39.

98 Commission Communication "Smart Regulation in the European Union", *supra* note 25, at p. 4.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press

ability versus learning, and different expectations about the appropriate actors to be involved in the process.

With its focus on the policy cycle, the new approach to evaluation seems inspired by a strong commitment to strengthening a solid evidence basis for European policy-making, whether expenditure or regulatory intervention. However, the challenges are so considerable that it may be that efforts focus on ensuring that *ex post* evaluation provides evidence for *ex ante* assessment of new action in relation to one particular objective, namely reducing the regulatory burden. The context in which the new approach to evaluation has developed, and the concrete actions taken so far, strongly confirm the likelihood of such a development.

https://doi.org/10.1017/S1867299X00004244 Published online by Cambridge University Press