

Impact assessment indicators

Measuring the quality of impact assessment

Claudio M. Radaelli & Anne C.M. Meuwese
Centre for European Governance, Department of Politics, University of Exeter

Abstract

Measures of regulatory quality are prominent in the academic discussion and have appeared in the policy debates on quality assurance, evaluation, and ex-post review of regulatory tools and institutions. In this chapter we focus on indicators of the quality of individual assessments and IA programmes. We introduce the concept of indicators and apply it to regulatory quality and IA. We also deal with an important distinction, between measuring quality as part of policy processes and measuring quality in a research context. There is no shortage of indicators. The key question for the future is how to integrate measures of regulatory quality with processes where policy-makers utilise them. This is as much a political question as it is a matter of proper scientific design.

Introduction

Regulatory quality is an elusive concept and the malleable framework that is Better Regulation does not do much to shed light on it. As an illustration of this, let us think of the following questions encountered regularly in the implementation of better regulation initiatives. Is the regulatory framework of high quality if it contains as little regulation as possible? Is 'good' regulation rather regulation that supports sustainable development? Or should we simply aim for procedures through which regulation is produced? Over the years, academics and policy practitioners have shed light on the characteristics of regulatory quality (Baldwin and Cave, recent book by Braithwaite, Mandelkern group) such as transparency, efficiency, access to legislation, participation, legitimacy, and responsiveness. These, however, are highly abstract concepts. Without measures it is not possible to evaluate whether a regulatory system is transparent or efficient. Measuring quality is notoriously difficult to realise because of (a) lack of agreement on standards in a country or, in the context of international organisations, across countries (b) intrinsic problems of operationalisation of concepts – it is daunting to imagine a set

of indicators of responsiveness of regulation; and (c) lack of useful data (see Chapter 6). Indicators present challenges, and when operationalisation is particularly challenging they should be supplemented by other measures, such as qualitative case studies, comparisons of processes, and participatory evaluation of better regulation policies. With these limitations, indicators are very useful to reduce ambiguity, make notions of quality explicit, and reveal preferences. When indicators are adopted by policy-makers, they contribute to the emergence of commitment.

The objective of this chapter is twofold:

1. to present an overview of the work that has been done on regulatory quality indicators and
2. to discuss how to design and use indicators for the measurement and evaluation of the quality of IA in both policy contexts and research contexts.

The structure of this chapter is as follows. First we explain what indicators are in a general sense and describe different types of indicators. Then we move on to the specific topic of indicators of regulatory quality, taking the reader through the growing literature in this field. Third, we pay special attention to the issue of how to design indicators. The chapter finishes by elaborating on the use of indicators in research and the use of indicators in public policy

What are indicators?

A useful definition if one wants to understand the function of indicators is given by the Joint Research Centre of the European Commission (2002):

‘[i]ndicators are pieces of information that summarise the characteristics of a system or highlight what is happening in a system’.

A second, more elaborate definition has been drawn up by the Tavistock Institute (2003):

‘An indicator can be defined as the measurement of an objective to be met, a resource mobilised, an effect obtained, a gauge of quality or a context variable. An indicator produces quantified information with a view to helping actors concerned with public interventions to communicate, negotiate or make decisions.’

In other words, indicators are a useful tool to draw conclusions on the performance or implementation of a principle, objective, concept or standard which in it-

self is impossible to measure. Indicators are operationalizations; if a measure captures the concept or principle or standard we have in mind we say that the indicator has internal validity, which means that ‘the instrument measures what it is supposed to measure’ (Patton 2002, p.14). The challenge, therefore, is to find valid measures of something other than the predefined principle, objective, concept or standard one is interested in, something that can be measured but can still reasonably be expected to give an indication of the performance of that principle, objective, concept or standard. Thus, indicators are one way of reducing complexity. For instance, instead of attempting to measure the ‘freedom of press’ one measures the proportion of quotes from opposition politicians in national newspapers. And to present some examples closer to the subject that concerns us: the number of pages in the Federal Register (Viscusi et al. 1995) or the number of regulatory agencies (Gatti 1981) have been used as a measure of the extent of regulation.¹ In the literature most relevant to the subject of impact assessment, ‘quality’ will often be the central concept, but this is not essential by any means, see for instance the indicators of social cohesion (Council of Europe 2005). Nor is the use of indicators restricted to the social sciences; indicators are also used to measure concepts like ‘biodiversity’.

Types of indicators

A first, basic typology distinguishes between individual indicators and composite indicators. Individual indicators are stand-alone measurements, varying from simple formats (typically ‘yes/no’) to more sophisticated ones. When statistical sophistication is used in order to capture a complex phenomenon using several indicators by aggregating their individual results, we speak of ‘composite indicators’.

A second important typology hinges on the method of measurement and the data used. The three types of indicators flowing from this distinction are set out below, each with an example from the IA context. This distinction is borrowed from a methodological guide on social cohesion indicators by the Council of Europe (2005).

1. quantitative and objective indicators (directly measurable, e.g. the number of IAs produced in a year);
2. qualitative and objective indicators (not quantifiable but based on objectively verifiable responses, e.g. whether a policy proposal has led to a legislative act being adopted);
3. qualitative and subjective indicators (perception-based; referring to an assessment or an opinion, e.g. whether an IA has sufficiently taken into account the results of a consultation).

¹ Examples taken from Radaelli & De Francesco 2007b.

A third typology is based on different dimensions of quality. Several labels are being applied; for instance Ladegaard (2007) distinguishes compliance tests, performance tests and function tests for impact assessment systems. In this chapter we will apply the following three types of indicators, which are closely linked to those three types of tests:

1. design indicators;
2. content-process indicators;
3. impact indicators.

Although these typologies and the distinctions they make are analytically important and useful to keep in mind when designing indicators, their relevance should not be overestimated; all distinguishing factors can become subject of debate. After all, composite indicators consist of individual indicators, although it is advisable to decide before designing the individual indicators whether aggregation is the final goal. Also, the difference between subjective and objective indicators is sometimes fluid. A question like ‘does your country have a better regulation policy?’ seems objective, but the answer depends among other things on one’s definition of ‘better regulation’. The distinction design/content-process/impact is not clear-cut either; they have to be used in conjunction. For instance, if indicators concerning the actual process and content also take into account elements that have been excluded from the design, the resulting aggregated picture is based on unfair assumptions.

Indicators of regulatory quality

The literature on indicators of regulatory quality can be divided into the following categories: economic literature, studies using surveys, studies using scorecards (see also chapter 6), and specific studies on indicators of regulatory quality. Both the World Bank and the OECD have been major sponsors of the debate among policy-makers and economists on regulatory quality indicators. Cooperation between each of these organisations and scholars has led to considerable progress in the development on indicators of regulatory quality.

Many of the comprehensive studies that have used indicators to measure regulatory quality have focussed on a specific area of regulation. An important example is the database on objective indicators of product market regulation developed within the context of the OECD, designed to produce one overarching indicator measuring the extent of regulation. (OECD 2006, Conway et al. 2005). The World Bank indicators cover the following regulatory sectors: employment protection law, bankruptcy laws, access to credit, property registration, protection of investors, and contract enforcement.

Other studies have produced indicators aiming to measure regulatory quality across sectors. One of the most ambitious projects is the World Bank set of

indicators on governance, which considers regulatory quality as one of the dimensions of governance and aims to carry out cross-country comparisons. The study that came out of this project (Kaufmann et al. 2005; 2003) aims to make objective use of subjective data (perceptions of a wide spectrum of regulatory stakeholders, including firms, experts or citizens contained in databases produced by international organisations, NGOs, and think tanks). Indicators are aggregated into an overall index ranking 209 countries according to the performance of their governance system ranking over five time periods (years: 1996, 1998, 2000, 2002, and 2004).

To the extent that the studies discussed in the foregoing look at regulation, they all tend to focus on the *quantity* of regulation. This very specific interpretation of quality is at least partly due to the bias inherent to the stakeholders whose perspectives (business, international economic organisations) have been chosen in those studies. The sensitivity of many systems of indicators to the preferences of stakeholders is a common problem, but should be seen as a direct consequence of the fact that indicators as a tool are usually (or at least potentially) linked to policy recommendations.

In a recent study the OECD has further expanded its work on indicators in the context of regulatory management, as developed by the OECD Programme on Regulatory Reform (Jacobzone et al. 2007). The study reiterates the difficulty of developing a set of mechanisms for assessing the likely impact of regulatory policies and for managing trade-offs – unlike in areas such as fiscal interventions where such mechanisms are commonplace. This explains why the OECD study takes an indirect approach and prioritises indicators that capture whether countries comply with good practice as identified by the OECD principles of good regulation.

A further strand of literature makes use of surveys. These studies often do only make use of indicators in a rather limited sense, i.e. that the information produced gives an indication of the *perception* of regulatory quality by business. The most important example here is the work by Pryor (2002a; 2002b), who has constructed an index of *laissez faire* by ranking countries according to their regulatory quality. However, since this index is based on selected indicators for certain markets (product markets, financial affairs etc.) his findings – as fully acknowledged to Pryor himself – are tied to a specific definition of quality, possibly in line with business views on the subject.

Finally, some have measured the quality of impact assessments specifically make use of scorecards. It is debatable whether the scorecard approach can be said to use indicators as the notion of quality often remains underdeveloped or implicit in these studies. Because many scorecards consist of a long list of individual and simple indicators their aggregated value is limited (cf. the discussion on weaknesses of composite measures above). Scorecards are based on assumptions that should be made explicit. Imagine that a scorecard is based on the value of 1 if an impact assessment monetises benefits, 0 otherwise. The assumption is the benefits should be always monetised, even if the data available are poor or extremely biased, and even if proportionality of analysis would suggest that it is not

efficient to spend too much time in trying to measure benefits in the context of the regulatory matter under discussion in the IA. It is also difficult to use scorecards to measure whether an IA was used by decision-makers or not – yet this is arguably the most important reason why IAs are prepared.

Thus, the simplest scorecards contain the assumptions that more analysis and more quantification is always a good thing. The first assumption is questionable when the scope of regulation is narrow and therefore investing considerable time and resources in ‘analysis’ dissipates taxpayers’ money. The second is problematic given the well-known limitations of cost-benefit analysis. Overall, basic scorecards are silent on the most important dimensions of quality of an IA, such as whether analysis was proportionate, the trade-offs presented by quantification, and the utilisation of analysis. Perhaps scorecards are useful in detecting ‘significant flaws’ in a system, but for a comprehensive measurement of quality, a more sophisticated measure is needed. Nonetheless, we present a brief overview here of scorecard studies on impact assessment in Europe, because these studies contain an important part of the limited data that is available on the subject so far.

Renda (2006) has used the scorecard approach to measure the quality of the extended IAs produced by the European Commission. His study shows two options that can be chosen. One can use the scorecard to measure how an institution complies with its own written guidance on IA or select exogenously the benchmark for quality. Renda opts for the latter, and consequently his scorecard looks for the presence of features of quality in IA carried out in year X that were not demanded by the Commission’s Guidelines in use in year X (Dunlop).

Lee and Kirkpatrick (2004) analysed a sample of six impact assessments completed by the European Commission in 2003 in order to assess the quality of both the process and the content. A similar study but one which used a yes/no format checklist instead of a scorecard was done by the Institute for European Environmental Policy in 2004 reviewing eight IAs from 2003 in order to determine whether sustainable development was included. This study made use of four dimensions: the written guidance, the operations of the IA systems, the IA report and the economic assessment in the IA. EU Impact Assessment has also been reviewed using a checklist approach by the Danish Environmental Assessment Institute (IMV 2006). Often, the reason for preferring a checklist over a scorecard is to minimise or even avoid the subjective interpretation that is involved in ‘scoring’. On the downside, checklists tend to underexploit the potential richness of the data.

Indicators of regulatory quality in the EU

A comprehensive list of indicators of regulatory quality defined in a broader sense has been produced as part of a study funded by the European Commission (DG Enterprise and Industry) in 2004. These indicators were designed mostly to be used by governments and administrations, who are also stakeholders to regulatory policy processes. The final results of the study were presented at a conference in

Brussels in January 2005. In its March 2005 Communication on Better Regulation for Growth and Jobs in the European Union, the Commission stated that it wanted to discuss the development of a common set of indicators and encourage their adoption by Member States to define targets and priorities for their better regulation programmes for the coming years (COM (2005) 97, 10). In the meantime a 'Group of High Level National Regulatory Experts' has been established to provide an interface between the Commission and the regulatory authorities in the Member States and to advise and contribute to the spread of best practices and the development of common indicators for better regulation. This idea continues to be pursued by the Secretariat General of the European Commission (Radaelli 2007). Although the Commission considers the indicators vital for the monitoring of the National Reform Plans under the Lisbon agenda, several Member States are hesitant on the issue.

DG Enterprise study, further elaborated in a research monograph (Radaelli and De Francesco 2007), proposes three systems of indicators, each representative of a certain phase in development of regulatory policy in a certain country:

- 1) Quality of the process: objective baseline indicators (measured by looking at written guidance and official documentation);
- 2) Internal evaluation: sophisticated real-world indicators;
- 3) External evaluation: (bridge between measurement of regulatory quality and the systematic evaluation of better regulation as public policy).

Obviously, if the guidelines do not perform well in this dimension, the automatic expectation is that micro-indicators of individual IAs show low quality. Except for random cases, bad guidance cannot lead to individual IAs of high quality. Good guidance however does not necessarily lead to good IAs, which implies that the first dimension identified by Radaelli and De Francesco should be seen as a filter: it is a logical pre-condition for the analysis of the other two dimensions. It is important to note that the Radaelli/De Francesco indicators are explicitly not aggregated, because they are designed for aggregation to be performed by the European Commission and national administrations, which should choose their own weightings to match the political purposes of the evaluation exercise.

How to design indicators

Defining quality

Indicators are linked to evaluation; indeed one property of good indicators is that they measure success or failure unambiguously. This implies that one has clarified what is understood as success before attending to the design of indicators. Another way to put it is that indicators have considerable potential for revealing preferences. When policy-makers or governments select indicators, they provide the clearest operationalization of what they mean by 'quality' or 'policy success'. Success in impact assessment can be defined in a variety of ways.

In EVIA we have defined quality in relation to the integration of different dimensions – in particular, sustainable development (SD) – in impact assessment. In turn, integration has been measured at different levels, that is, design, content, and process of IA, with particular emphasis on participatory processes. Survey indicators were deemed an appropriate tool for this task. We have made the disclaimer that our notion of quality does not apply to every individual IA, as IAs exist for which the SD dimension is not relevant. Besides, the distinction between social, environmental, and economic dimensions is somewhat artificial. There is an economic dimension to social policy, as well as an environmental dimension to economic policy goals. Furthermore, environmental policy objectives have obvious economic implications, and so on. However, the general point about the integration of different dimensions in the design of IA systems across countries and in the performance of specific IAs is relevant. For instance, the consideration (or lack of) of external implications of domestic rules (for example on developing countries) and the endogeneity of the preferences of future generations are important ways in which an IA can be integrated or not.

Selecting indicators

Two basic methods of selecting indicators can be identified:

1. Bottom-up: to start with a long list of possible indicators and then narrow it down as the research advances, either (preferably) on the basis of meta-criteria or on the basis of practical experience.
2. Top-down: to start with defining the principle, objective, concept, or standard which one wants to measure, breaking it down into different dimensions.

For each of these dimensions appropriate indicators should then be composed. An example of the application of the top-down approach can run as follows: the various phases of the regulatory process are taken as a starting point, subsequently the literature on each respective phase for common standards for quality is reviewed, and on this basis, design ways to measure the performance against these standards are determined. Regardless of which selection method is followed, it is important to keep in mind that an indicator is always a measure of something other than the principle, objective, concept, or standard of primary interest - which are not measurable as such. Consequently, internal validity has to be properly checked. The design of indicators can only start after the choice on the purpose and the method for the policy appraisal has been made. The appropriate way to design indicators also depends on the type of indicators one is looking for. If an aggregated result is important, the top-down method of composing indicators is probably more appropriate. For individual indicators the bottom-up method can be fruitful, because it may yield more creative indicators. However if a decision to aggregate after all is subsequently made, the interlinkages between indicators and their proportionate weights as well as their relationship to the larger principle, objective, concept, or standard one wants to measure are likely to be unclear.

Trade-offs in indicators design

The Council of Europe in a report on indicators of social cohesion proposed some general quality criteria for indicators (Council of Europe 2005). According to this report indicators should:

- be representative with regard to the question to which they relate;
- provide a wealth of information;
- be unambiguous;
- be based on clear and widely shared interpretation;
- not be excessively expensive

Even if this is a good checklist to keep in mind when designing indicators, there will inevitably be trade-offs between these criteria. It is likely that many indicators which provide 'a wealth of information' will be the most expensive to collect. Paradoxically, the overarching trade-off problem in designing indicators is the one between reliability and usefulness. Since indicators try to bridge the gap between a seemingly immeasurable concept that often includes a value judgment and a vast amount of descriptive data, the results they produce should not be treated as hard evidence. The bigger the gap, the more interpretative steps will be needed and the less reliable the findings will be. At the same time the most useful indicators are bound to be those bridging the largest gaps.

For example, when designing the more ambitious 'impact indicators' the trade-off between reliability and usefulness applies in the following way: it is ex-

tremely difficult to provide robust indicators of the causal relationship between a specific feature of an IA and economic variables. The causal link between IA of a given rule, the rule eventually adopted, and the impact of rule on the socio-economic system may be interrupted at several points by intervening variables. Another way of looking at impact is to consider how IA changes the attitudes of the regulators about regulation.

Using indicators in research contexts

Just like all evaluation tools, indicators have advantages and limitations. First of all, it should be stated that indicators are but one way of evaluating policies and should ideally be complemented by other methods such as case studies, qualitative reviews and formative appraisals. Secondly, the margins of error can be substantial, not only when subjective indicators are used but also in the case of objective indicators. This is not only because the data can be erroneous, but also because the extent to which the data give an accurate indication of the performance on the principle, objective, concept or standard which is the final object of the evaluation at hand is often doubtful. Furthermore, objective indicators sometimes make use of data collections that are based on subjective insights. It should also be noted that wide margins of error are a well-known feature of all methods to measure governance-related concepts (Kaufmann et al. 2005). A third problem is that the fierce debate on whether indicators should be used to produce an aggregate index remains unresolved. Can a list of individual indicators still capture reality? If the answer is to be found in designing composite indicators, to what extent is the weighting that is involved arbitrary (Joint Research Centre 2004)?

An example of the use of indicators for research purposes is described below. We start from some four dimensions. They are institutionalisation, purpose, sophistication and accessibility. Each dimension can be represented by a continuum from low to high.

Table 1a

The dimension of 'institutionalisation'

	Examples of indications of low institutionalisation		Examples of indications of high institutionalisation
Degree of institutionalisation	IA requirements on departmental level only No mechanisms for quality control No training	↔	IA requirements enforced by law Legal mechanisms for quality control IA training is part of the regular curriculum for civil

No central support	servants Central support units
--------------------	-----------------------------------

Table 1b

The dimension of 'purpose'

	Examples of indications of narrow purpose		Examples of indications of broad purpose
Purpose	Strict and limited demarcation of impacts to be assessed No requirement to define objectives and problems clearly in impact assessment Overall aims not defined in a governmental agenda or policy papers A priori focus on one type of costs/benefits	↔	Encouragement to perform IA on side effects, long term impacts, trade offs, inter-linkages, Requirement to set objectives beyond the purpose of an individual proposal Overall aims of the system laid down in a governmental agenda Focus on the overall cost/benefit ratio

Table 1c

The dimension of 'sophistication' of economic analysis

	Examples of indications of low sophistication		Examples of indications of high sophistication
Sophistication	Checklist approach No consideration of trade-offs IAs mainly justify the proposal Data taken at face value No external expertise used Status quo is not analysed No consideration of scientific literature	↔	Causal analysis (of how A causes B) Use of models to examine trade-offs Exploration of potential side effects Critical assessment of quality of data and sensitivity analysis Inclusion of external expertise Elaborated baseline scenario Peer review of the analysis included in the IA

Table 1d

The dimension of 'accessibility'

	Examples of indications of a closed system		Examples of indications of openness
Accessibility	IA results not public	↔	Systematic publication of IAs on websites

Using indicators in policy contexts

Policy-making occurs through interaction. Interaction depends on resource interdependence, including information. The main added value of indicators lies in their potential to add a new kind of information (Radaelli & De Francesco 2007a). However, to use indicators in a policy evaluation study is one thing; to use them in policy processes is quite another. Although in the definition by the Tavistock Institute (see above), actual policy use is seen as a core feature ‘helping actors concerned with public interventions to communicate, negotiate or make decisions’, it is also their most thorny side. Using indicators in actual processes of regulatory policy development implies that political conclusions are tied to the outcomes of the indicators. Translating this issue to the IA context specifically: the more formalised a set of indicators is and the more open to evaluation, the more their use can amount to controlling the development of the IA system. This mechanism could explain the considerable reluctance to adopt them on the part of those working with the IA system on a daily basis. Thus, the more useful a set of indicators potentially is, the harder it will be to agree on the design, especially when it includes an aggregate index. Here the problem that quality is not neutral vis-à-vis different stakeholders in IA (Radaelli 2005) presents itself again. Since indicators are tied to specific notions of quality, agreement on indicators is a political milestone. And even if agreement could be reached – for instance on a list of rather loosely composed indicators – problems are likely to arise with regards to the use of the information produced by the individual indicators. Here it is important to have a forum (like annual reviews, or hearings in Parliament, or more sophisticated open coordination platforms) where the information produced by indicators for monitoring purposes is used to kick-off discussion and learning-oriented usages of information.²

The recent experience with implementing indicators in the development of regulatory policy at the EU level tentatively confirms that tying consequences to a set of indicators represents the greatest challenge. The policy discussion on using indicators in the EU context has largely been of a technical nature and has lacked concrete commitments. Perhaps this is unsurprising given the diffusion of principles of better regulation in the EU and the great variation in implementation of IA. Still, if policy-makers are committed to regulatory quality, they should make use of measures to show results to their stakeholders. According to the latest accounts some governments ‘are “reluctant although not opposed” to make the step of adopting common indicators’ (Radaelli 2007). Political hesitation notwith-

² On the relationship between monitoring and learning see Radaelli and De Francesco (2007a) who in turn draw on the path-breaking intuition of Charles Sabel.

standing, statements on the desirability of using indicators in EU regulatory policy are plentiful. There is for instance the Four Presidencies Initiative of January 2004 declaring that ‘the Commission should propose indicators to measure progress with regulatory quality and reform at European and Member State level for activation through the Open Method of Coordination and for application in impact assessments’. The presidencies also stated that ‘[p]rogress can best be measured if such indicators are quantitative’. Concretely, ‘the administrative burden for companies’ is proposed as a useful indicator, although it remains unclear of what quality aspect it would give an indication. Targets for administrative burden reduction and systems of indicators of regulatory quality should be kept analytically separate: one cannot achieve quality assurance goals by using the former and ignoring the latter.

The High Level Group of National Regulatory Experts has put indicators on the policy agenda, but they were never formally adopted – only discussed. The informal body consisting of civil servants coordinating EU regulatory policy at the member state level known as ‘Directors of Better Regulation’ (DBR), which overlaps greatly with the High Level Group in terms of its members, has also held sessions discussing indicators of regulatory quality in more general terms (Radaelli 2006). The results of a 2007 survey held among these European experts on Better Regulation in the context of the Intune research project (www.intune.it) also suggests that the adoption of indicators is not felt as a priority in this policy community. Only 20% of the respondents chose regulatory indicators as one of the three major items that should feature on the future BR agenda.

The Commission communication on Better Regulation for Growth and Jobs in the European Union from 2005 contains the following wording on indicators:

- ‘The Commission intends to discuss in this group the development of a coherent set of common indicators to monitor progress as regards the quality of the regulatory environment both at EU level and in the Member States themselves, as a basis for benchmarking. The Commission will encourage Member States to adopt such indicators to define targets and priorities for their better regulation programmes for the coming years in their national Lisbon programmes.’ (European Commission 2005, p. 10)

Tellingly, a later policy document of the same agenda-setting stature does not mention indicators at all (European Commission 2006). In the meantime, the Council has agreed on targets for the reduction of administrative burdens in March 2007.

It is the use of indicators rather than their design, which is pivotal to their success as aid in policy processes. First of all, indicators are more likely to be used if they capture variables on which policy-makers can intervene – a large number of variables affecting the business and social environment are not controllable by governments. Secondly, the types of indicators that are likely to win agreement

among all political actors are also likely to be the ones that are amenable to various uses and interpretations. To illustrate this point we borrow from the indicators debate in the audit sector: Scott (2003) observes that even if the use of non-financial performance indicators has become fashion among auditing institutions, these are still interpreted through the lens of the traditional pre-occupation with narrow financial controls. Similarly, actors using indicators in policy processes are bound to do so with their personal or institutional notions of 'better regulation' and 'best practice impact assessment' in mind.

Apart from the bias problem, and regardless of the method of selecting indicators, one either needs meta-criteria (method 1) or well-defined quality dimensions (method 2) to enable indicators to actually indicate something beyond their direct objects of measurement, instead of reducing their outcomes to fragmented pieces of information. For instance, it is impossible to design an indicator measuring whether IA fosters 'subsidiarity-wise' regulation without a clear concept of what is meant by subsidiarity in the first place. The only way to develop a meaningful system of indicators is to develop a consensus on what the EU wants out of impact assessment and better regulation more generally.

The use of a set of indicators that is integrated in the policy process could help the Commission to monitor IA quality and report on the progress in a systematic way so as to become more accountable to stakeholders. Furthermore, indicators could help the Member States to make progress in a common direction, provided that this common direction has been defined in advance. Here we are back to the core of the problem: the definition of goals is linked to measuring their attainment, the stakes are raised and agreement becomes more difficult to attain. However, if indicators are seen by governments as a diagnostic tool to assist in identifying priority areas for improving their regulatory policies (Jacobzone 2007), designing common indicators becomes a real possibility. The Directors of Better Regulation (DBR) can be the forum where data are collected and technical discussions on indicators take place.

Some big questions in the debate are: to what extent does quantification assist in the integration of different dimensions? What is the best balance between quantification and monetization? Why is quantification so difficult in Europe – as opposed to the USA? In order to start addressing this question, measures are indispensable. An implication is that we have to be clear about the implications of our measures: if an indicator shows that quantification is difficult, is this a good or a bad thing in relation to our notion of quality? If indicators show that regulators use a model instead of a framework to assess environmental policy issues, is this an improvement in terms of quality or not? What is the role of trade-offs between desirable dimensions of integration and how can this be measured via indicators? These are difficult questions, but they cannot be ignored, because they lead to a strong notion of quality, that goes beyond the accuracy of cost and benefits estimates.

Another change of perspective that could forge a breakthrough in the impasse on using indicators in a policy context is to distinguish between different levels of indicators. Going back to the three-tiered system of indicators of Radaelli

and De Francesco, the key is to ensure the set of indicators fits with the stage of implementation of a state's impact assessment system. It is crucial to find the right timing for moving from macro indicators on guidance documents, to micro (at the level of individual IAs) indicators suitable for internal evaluation, and then to pluralistic external evaluation³ to consolidate better regulation as public policy. In order to achieve the latter, strategic and operational management, specific structures, and dedicated tools are required.

To date there is only one real-life example of the third type of evaluations and even there the indicators approach was not taken explicitly: the evaluation of the European Commission's impact assessment system by The Evaluation Partnership (2007). A peer review of the better regulation policies in the EU15 member states is under way at the OECD (funded by the European Commission) but it is not clear to what extent the peer reviewers will make use of indicators. Those member states which so far only have basic approaches to regulatory quality and impact assessment programmes in pilot stages could be content with the baseline indicators of the first level. They could already adopt some of the more simple indicators of level two, which are on the whole more suitable for member states with experience in consultation, simplification, and the assessment of administrative burdens if not with impact assessment per se.

Concluding remarks

As this chapter has shown, there is no shortage of possible indicators to be used. In this chapter we have introduced different types of indicators, and discussed their trade-offs. One challenging issue is how to fit regulatory indicators to the demands of (academic) research and, most crucially, the institutional and political settings of the environment in which information is produced and exchanged, taking into account the given that information can be used to mislead rather than inform the public or not used at all.

Indicators present their own set of challenges, but they provide opportunities too. They are a tool to reveal the preferences of policy-makers. If properly supported by operational and strategic management, they can also link knowledge and utilization for policy purposes. In the long-term, they are a pre-requisite for constant improvement and convergence of better regulation policies in Europe. Finally, and perhaps more importantly in the long term, regulatory indicators are a transparent way through which better regulators can communicate with citizens and firms, and support a wide social debate on the results achieved by better regulation policies. As such, their potential for regulatory accountability is considerable.

³ Pluralistic evaluation refers to a situation in which different governmental and non-governmental bodies engage in policy evaluation from different perspectives. See Schwartz, R. and J. Mayne (Eds) (2005). *Quality Matters: Seeking Confidence in Evaluating, Auditing and Performance reporting*. New Brunswick/London: Transaction Publishers, p. 25.

Bibliography

- Conway, P., V. Janod, and G. Nicoletti (2005). 'Product market regulation in OECD countries: 1998 to 2003'. Economic Department Working Papers, No. 419, ECO/WKP(2005)6. Paris: OECD.
- Council of Europe (2005). Concerted development of social cohesion indicators: Methodological guide. Belgium.
- Dunlop, C.A. (2006). 'Impact Assessment in the EU: The State of the Art and the Art of the State' (Renda), ERA - Forum Number 1.
- European Commission (2005). Communication from the Commission to the Council and the European Parliament - Better Regulation for Growth and Jobs in the European Union, COM(2005) 97 final.
- European Commission (2006). Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions - A strategic review of better regulation in the European Union, COM(2006) 689 final.
- IEEP Institute for European Environmental Policy (2004). 'Sustainable development in the European Commission's integrated impact assessment for 2003', Final report, April 2004.
- IMV Environmental Assessment Institute (2006). 'Getting Proportions Right. How far should EU impact assessments go?'. Copenhagen, April 2006
- Jacobzone, S., C. Choi and C. Miguet (2007). 'Indicators of Regulatory Management Systems', OECD Working Papers on Public Governance, 2007/4, OECD Publishing.
- Joint Research Centre of the European Commission (2002). 'State-of-the-art report on current methodologies and practices for composite indicator development'. Paper prepared by the Applied Statistics Group. June 2002. Ispra, Italy: Institute for the Protection and Security of the Citizen Technological and Economic Risk Management.
- Joint Research Centre of the European Commission (2004). 'Composite Indicators - The Controversy and the way forward', a presentation at the Statistics, Knowledge and Policy: OECD World Forum on Key Indicators by A. Saltelli, M. Nardo, M. Saisana and S. Tarantola. Palermo November 2004
- Kaufmann, D., A. Kraay and M. Mastruzzi (2003). 'Governance matters III: Governance indicators for 1996-2002'. World Bank Policy Research Working Paper, No. 3106. June 30, www.worldbank.org/wbi/governance/govdata.
- Kaufmann, D., A. Kraay, and M. Mastruzzi (2005). 'Governance matters IV: Governance indicators for 1996-2004'. World Bank Policy Research Working Paper, No. 3630, May 9, www.worldbank.org/wbi/governance/govdata.
- Ladegaard, P. (2007). Measuring RIA quality and performance. In C. Kirkpatrick and D. Parker (Ed.), Regulatory Impact Assessment: Towards Better Regulation?. Cheltenham: Edward Elgar, pp. 56-71.
- Lee, N. and C. Kirkpatrick (2004). 'A pilot Study on the Quality of European Commission Extended Impact Assessment'. Impact Assessment Research Centre, Institute for Development Policy and Management, University of Manchester, 21 June.
- Organisation for Economic Co-operation and Development (OECD) (2006). Economic Policy Reforms: Going for Growth 2006. Paris
- Patton, M.C. (2002). Qualitative Research & Evaluation Methods, 3rd edition, Thousand Oaks: Sage Publications.
- Pryor, F. L. (2002a). The Future of U.S. Capitalism. Cambridge: Cambridge University Press.
- Pryor, F. L. (2002b). 'Quantitative notes on the extent of governmental regulations in various OECD nations', International Journal of Industrial Organization, 20: 693-714.

- Radaelli, C.M. (2005). 'Diffusion without convergence: How political context shapes the adoption of regulatory impact assessment' *Journal of European Public Policy*, 12(5), 924-943.
- Radaelli, C. M. (2007). Whither better regulation for the Lisbon agenda?, *Journal of European Public Policy*, 14(2): 190-207.
- Radaelli, C.M. and F. De Francesco (2007a). *Regulatory Quality in Europe: Concepts, Measures, and Policy processes*. Manchester: Manchester University Press.
- Radaelli, C.M. and F. De Francesco (2007b). 'Indicators of regulatory quality'. In Colin Kirkpatrick and David Parker (Eds.), *Regulatory Impact Assessment: Towards better regulation?*, Cheltenham: Edward Elgar, pp. 36-55.
- Renda, A. (2006). *Impact Assessment in the EU. The State of the Art and the Art of the State*. Brussels: Centre for European Policy Studies.
- Schwartz, R. and J. Mayne (Eds) (2005). *Quality Matters: Seeking Confidence in Evaluating, Auditing and Performance reporting*. New Brunswick/London: Transaction Publishers.
- Scott, C. (2003). Speaking Softly Without Big Sticks: Meta-Regulation and Public Sector Audit, *Law & Policy*, 25(3):203-20.
- Tavistock Institute (2003). *The Evaluation of Socio-Economic Development*, December 2003, www.evaled.info.
- Viscusi, K., J. M. Vernon, and J. E. Harrington (1995), *Economics of Regulation and Antitrust*, Cambridge MA, MIT Press.