DESPERATELY SEEKING
REGULATORY IMPACT ASSESSMENTS:
DIARY OF A REFLECTIVE RESEARCHER

Claudio M. Radaelli

Director, Centre for Regulatory Governance,
Professor of Political Science and Jean Monnet Chair,
Department of Politics, Rennes Drive
University of Exeter, Exeter EX4 4RJ, United Kingdom

C.Radaelli@ex.ac.uk

Draft – not for quotation

PAPER DELIVERED TO THE WORKSHOP OF THE
EUROPEAN NETWORK FOR BETTER REGULATION
ON “BETTER REGULATION AND THE POLICY PROCESS”

LISBON, 5 JULY 2007

The support of the Economic and Social Research Council,
research grant on “Regulatory Impact Assessment in
Comparative Perspective” No. RES-000-23-1284 is gratefully
acknowledged
DESPERATELY SEEKING REGULATORY IMPACT ASSESSMENTS: DIARY OF A REFLECTIVE RESEARCHER

Claudio M. Radaelli

Abstract
This paper provides some reflective thoughts on current research on regulatory impact assessment in Europe. The narrative draws on the author’s fieldwork diaries. It starts from the methodological problems of identifying and analyzing regulatory impact assessment in the UK, Denmark, the Netherlands, and Sweden. The root of the problem, however, is not methodological. It is about research design (specifically, concept formation and classification), the relationship between supply and demand of research, and, ultimately, who defines the objects of inquiry. The conclusions draw lessons on how to improve on research design and present the basic components of a more balanced relationship between the client and the researcher.

Key-words: Regulatory impact assessment, comparative public policy, qualitative methods, research design, knowledge utilization, fieldwork

We frequently argue that apples and pears are ‘incomparable’; but the counterargument inevitably is: How do we know unless we compare them?


1. DIARIES, COMPARISON, CONCEPTS, AND CUMULATIVE RESEARCH

In this paper I address the European Network for Better Regulation (ENBR) by presenting my reflections on another, smaller project on impact assessment. The main aim is to anchor my thoughts to something I know very well and ask you, the audience, if the issues aired in this paper resonate with your ENBR experience.

As many of you know, I am directly involved in ENBR - Anne Meuwese at the Centre for Regulatory Governance, University of Exeter, mans the helpdesk; I coordinated the high level group of experts who designed the handbook for data collection we use in ENBR. However, in order to maintain some distance from the ENBR project, I think it is better to approach a common problem of research and knowledge utilization by looking at the findings of another project.
The general problem we face as better regulation researchers has three dimensions, namely (a) concept formation, (b) the relationship between policymakers and research communities, and (c) the nature of our dependent variable, that is, what we really want to explain. There is also a fourth methodological component cutting across the other three, but, although I will use the language of research design and research methods extensively, the main root of the problem is not methodological.

To sum up then, this paper provides some reflective thoughts arising out of another project with a view to discussing whether similar problems are encountered in ENBR. Should the answer be positive, Anne Meuwese and I would like to come back to ENBR with a co-authored paper specifically targeting ENBR problems. In this introductory Section I present how I came to elaborate this paper by using my fieldwork diary, and explain how the problem has a methodological component - yet the most important dimensions are (a), (b), and (c) as illustrated above. Section 2 tells the story of what happens when one moves from abstract notions about better regulation and impact assessment and gets into the first steps of comparative qualitative fieldwork. Section 3 presents my reflections on the fieldwork carried out in the UK. Section 4 turns to the Netherlands, and shows how I was surprised not to find the regulatory impact assessments I meant to examine. Section 5 reports on how I became increasingly puzzled during my visit to Denmark. Section 6 explains the difficulty encountered in classification based on the fieldwork in Sweden. Section 7 draws conclusions and provides the questions I would like to discuss with my ENBR colleagues.

Most papers published in social sciences journals tell a story – often a causal story. The latter can be told with the aid of comparisons, case studies, econometrics, facts and figures, descriptions, political metaphors, and policy narratives. This paper tells its own story by using a diary. Indeed, it is based on segments of a research diary written at the end of key stages of fieldwork. It draws on the notes one writes to make sense of a period of fieldwork in a given country. Diaries, personal stories, and reflective journals are quite common in disciplines such as biology and educational sciences (Leshem & Trafford, 2006). Usually, the researcher examines the diaries of other individuals (students, patients in health care, etc.) but here I am thinking of ethnographic travelogues that blend the evidence collected during fieldwork with emotions such as surprise, frustration, disappointment, and excitement (Tedlock, 2000: 463). The point about the ‘reflective’ nature of the researcher in the title of this paper is a tribute to the well-known work of Donald A. Schön (Schon, 1983).¹

¹ The story in short: in his Reflective Practitioner, Schön contrasted the abstract, hyper-rationalist models that teachers, scientists, architects and others learn in school-universities with what they do in practice, and the problems they encounter in their professional experience. The gap is due to the neglect of the basic fact that problem setting in the real world always take place in a specific socio-institutional context, characterized by uncertainty, ambiguity and value conflict.
Here I do not wish to contribute to scholarly work on reflective diaries. Neither do I want to present them extensively and talk about ‘why did I write this at this point of my work’? The reference to diaries is more indirect, perhaps oblique. Diaries are not the main subject here. They are the starting point that enables me to tell you a story. I will use the present tense to enter a real ‘diary’ mode and give the flavor of the sense of surprise I experimented day by day.

But what is the story then? The story, as mentioned, is about how do we define impact assessment, how do we get to this definition and in relation to what kind of policy-making world ‘out there’, and, partly correlated with the two questions, how do we get to choose to score RIAs instead of looking at their emergence, types, and the role of the institutional-administrative context in the implementation of abstract notions such as better regulation.

RIA is a form of policy appraisal that enables legislators and regulators in general to appraise the impact of regulatory proposals by using several techniques, including the economic analysis of costs and benefits. We will come to the definition in a moment. At this stage, we simply note that there has been a wide diffusion of this tool in the context of policies to increase the capacity of governments to produce high quality regulation. These policies are often referred to as ‘better regulation’. The OECD has played a key role in the 1990s in the diffusion of better regulation and RIA. With the OECD ministerial declaration on regulatory quality (OECD, 1995), the Paris-based organization provided the first international standards in this policy area, endorsed at the highest possible level by its Member States. The regulatory review produced by the OECD have provided a single, common template for the examination of progress with tools like RIAs across a large number of OECD Member States.

Over the last ten years or so, the European Union (EU) has also emerged on the map of better regulation, both in the sense of uploading (the most active governments and Presidencies pushed for the adoption of their standards of better regulation as common EU goals) and downloading (the Commission has been extremely active in asking the better regulation laggards to make progress with the introduction of RIA, regulatory quality indicators, and initiatives to combat red tape).2

Unsurprisingly then, a member of our ENBR consortium has observed that RIA has now become a global norm (Jacobs, 2006). This global diffusion is explained by the potential of this tool: perhaps, in the not-so-distant future, RIA will do for the analysis of regulatory costs and benefits what finance bills and budgets do for public expenditure and taxation – a point that Scott Jacobs often makes at our ENBR meetings. Arguably, impact assessment can provide the spine of more sophisticated approaches to regulatory management, possibly in the form of

2 The terms ‘uploading’ and ‘downloading’ are frequently used in the literature on the Europeanization of public policy (Börzel & Risse, 2003).
regulatory agendas (Doern, 2007). Jacobs notes that 23 out of 30 OECD countries have established ‘formal policies mandating the use of RIA in domestic policy-making’ (Jacobs, 2006: 5). The OECD has established a RIA observatory to keep track of who does what in this area.

The diffusion of this tool of regulatory governance has induced researchers to move on from the previous, single country study - usually the USA (Coglianese, 2002; McGarity, 1991) - and matched comparisons (typically the USA and the UK (Baldwin & Veljanovski, 1984) or more recently the US and the EU (Cecot, Hahn, & Renda, 2007) to a large set of countries. In Europe, better regulation and RIA have become a political priority in the context of the Lisbon agenda for growth and jobs in the European Union (Radaelli, 2007). According to what governments report to Brussels, the vast majority of the EU-27 countries have introduced RIA or pilot projects. The European Commission is funding research to provide datasets of impact assessment practice in the Member States.

Briefly, there is strong demand for comparative research on impact assessment. The supply (academic researchers and policy institutes) has followed suit. The total number of comparative studies on RIA has increased dramatically. Problems of comparative analysis are well-known. In emerging fields of research, one perhaps should not be excessively concerned about these teething problems, and adopt some forms of crude operationalisation of concepts, rely on proxies, and make use of data that are not the ones we need but the ones we happen to have, and so on. There is more wisdom than arbitrariness in this approach: one starts with some measures that, crude as they may be, enable some empirical control of our conjectures. Other more sophisticated researchers – the argument goes on – will in the not-so-distant future criticize our measures, collect better data, and refine both measurement and theory testing. This is the beauty of the cumulative approach in the social sciences.

However, at some critical stages in the project, I have discovered the limits of this particular approach at my own cost – in terms of puzzlement, confusion, and blurred concepts. In the case of RIA at least, we (that is, the RIA researchers) need to pause for a moment and reflect on what the concepts really mean when they (and us researchers with too) travel across countries.

RIA is a hard nut to crack in terms of systematic comparison. This incredibly elusive object of analysis can lead one to produce negative results: we cannot prove much if there is no real comparison or if we compare apples with oranges. One can decide to throw these results in the ‘too hard to think’ bin: negative results cannot be published. Recently, however, there has been interest in negative results, as shown by the arrival of dedicated journals (Leher, Lesche, Ihachimi, Vasiliu, & Weiffen, 2007). More importantly still, one can learn a lot from the failure of research strategies by asking the question ‘what is it that it did not work during the fieldwork?’ To be more specific: one can learn whether the strategy does not work because apples are compared with oranges, and become
more aware of the differences between items in the same species and between species and their genus (Sartori, 1970). As shown by Brady and Collier, concept formation logically precedes measurement (Brady & Collier, 2004). These learning points can then be used to reformulate concepts, to provide more finely-grained toothcombs to tame empirical evidence, to clarify who defines what in the context of demand and supply of research, and ultimately to make (better) sense of a complex reality. Thus, there is no contradiction with the aim of contributing to cumulative research in political science.

2. PROBLEM, RESEARCH QUESTION, MEASUREMENT: BUSINESS AS USUAL?

Let me spend a few words on the project I mentioned. It is called Regulatory Impact Assessment in Comparative Perspective and it is funded by the ESRC, the Economic and Research Council of the UK (http://www.esrc.ac.uk). The project is comparative and empirical (as opposed to projects on individual countries and-or purely theoretical investigations). It is carried out by the author of this paper, together with a senior research assistant. So it looks like a fairly simply enterprise. The two of us know each other well. We have published a book together on regulatory indicators and have carried out other projects, including a book on RIA in Italian in 2001 (Radaelli, 2001). Between the two of us, there is no inter-subjective problem of meaning, interpretation, ambiguity. When we say ‘RIA’ we know what we are talking about – well, perhaps, we think we know... When we say ‘better regulation’ we can discuss a bit more, but in the end we know how the quality of this policy should be measured. However, the fact that we know each other well makes wrong assumptions more difficult to be detected. I’ll come to the issue of assumptions about the world out there in a moment.

The project covers seven political systems, that is Canada, Denmark, the Netherlands, Sweden, the UK, USA and the European Union. The research question is about the political changes brought about by regulatory impact assessment (RIA)? The focus is obviously on RIA, but we also situate this tool in the wider context of better regulation policies.

In the first part of the project we explained how better regulation and RIA are evolving in Europe. A large literature review assists us in formulating some initial hypotheses on the interplay between impact assessment and different trajectories of regulatory governance. We can therefore say that some elements of RIA are most likely to be associated to this broad political direction of regulatory governance rather than this other direction. We can also say that RIA can be used to foster one model of governance (such as control, learning, or symbolic politics) or another. Why this happens we do not know yet. Finally, we set out to develop our measures. With this background information in mind and
some early drafts of papers in our PCs, we are ready for fieldwork in the countries selected for this study.

Fieldwork in the countries covered by the project is informed by some assumptions. The first is that a definition of RIA exists – indeed in the context of ENBR we came up with a fairly standard definition of RIA (based on previous OECD definitions), portrayed in box 1 (ENBR, 2006).

<table>
<thead>
<tr>
<th>Box 1 – DEFINITION OF RIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulatory Impact Assessment is:</td>
</tr>
</tbody>
</table>

1) A systematic, mandatory, and consistent assessment of aspects of social, economic, or environmental impacts such as benefits and/or costs;
2) affecting interests external to the government;
3) of proposed regulations and other kinds of legal and policy instruments;
4) to i) inform policy decisions before a regulation, legal instrument, or policy is adopted; or ii) assess external impacts of regulatory and administrative practices; or iii) assess the accuracy of an earlier assessment.

[Source: ENBR, 2007]

The existence of a definition is important: it means that we know what we are looking for. International organizations like the OECD and the European Commission have established similar definitions (as mentioned, the ENBR definition draws on previous OECD work) and have come to the conclusion that RIA is pretty much established in Europe, especially in the countries selected for our project.

On the basis of this definition, we have selected countries that have a long-term experience of RIA (roughly, a decade), controlling for geographical distribution (the two sides of the Atlantic), complexity of governance (multi-level and unitary system), degree of delegation of regulatory power (high in Sweden, where the central government is comparatively speaking quite small, low in the Netherlands). We need a sample of countries with solid experience of RIA – this is the argument in our research design - because we want to examine the effects of this tool. Ten years is a pretty good time-span to check whether innovations, in our case RIA, have bedded in, perhaps mutating over time. If there are no changes whatsoever brought about by RIA in these countries, the probability of
finding political changes in countries with limited experience of RIA is low. So, to recap, we start with the assumption that we know what RIA (the ‘cause’) is, hence we can examine its effects.

The second assumption is that fieldwork can be organized around micro and macro-analysis. Micro analysis revolves around the collection of samples of RIAs in the different countries – across sectors, in agencies and central government, and so on. Macro analysis entails interviews with samples of policy-makers, employers’ associations, one or two academic experts, and research institutes in each country. We perform desk research before the interviews, and then we go back to the desk and read more when some interesting aspects of a political system crop up in the interviews. This is normal practice.

Comparison is an essential component of the project. For this reason, we design a standardized questionnaire. Although some questions are very open and some are semi-structured, a good deal of the questionnaire is based on scales and relatively closed answers. The reasoning is the following. In comparisons of two or three countries, it is still possible to get into explanations of how causality works within a given institutional and cultural sector and leave the grid for comparison very open: first we want to understand what RIA really “means” in countries A and B, then we will compare. Interpretation and dedication to the object of empirical analysis are the supreme principles. No-one wants to fly too high with abstract comparisons.

In our case, however, we have seven systems to control for. Too many cases for indulging too much in idiosyncratic interpretations - with the risk of concluding that ‘the Dutch RIAs have these consequences because they are made in the Netherlands’. We have too much interest in causality (does RIA, the cause, produce this change or that change in this country, and why not in the others?) for not trying to compare systematically across a set of relatively broad, macro independent variables. Hence questions that can travel across countries abound in our questionnaire. Note: we make the assumption that definitions and questions do indeed travel across countries, but we have not tested it. And with these assumptions in mind, off we go into the fieldwork mode.

3. CONCEPTS TRAVEL, SO DO RESEARCHERS: FIELDWORK STARTS IN THE UK

The first country selected for fieldwork in the UK. The idea is that it is better to play at home before playing away the other matches. We have observed the UK RIA scene since 2000. The two of us have talked to a great number of people in the cabinet office in the last three years. I have advised the National Audit Office on the topic of impact assessment.
The early set of interviews, however, is disappointing. The cabinet office has more than seventy people in a central unit called Better Regulation Executive. Yet most of them do not deal with RIA, although this is the flagship tool of better regulation. To avoid problems, we select people with strategic responsibility for better regulation, have spoken about RIA at several events, and, in brief, are quite knowledgeable. Some of them know us very well. We also decide to speak to people in the Better Regulation Commission, a sort of arm’s-length body that challenges the way the government implements its better regulation goals. We also arrange meetings in key departments, especially those who have a reputation for being quite good at performing RIAs.

The surprise is great when in interviews I ask the question ‘can you give us some examples of good RIAs’? The question – we reasoned when we designed the questionnaire – is not so much about which RIAs are selected by the interviewee, but a way of probing her on what makes these RIAs good. Yet there is often silence after this question. True, some well-organized civil servants have prepared a list of RIAs that, I cite the title of the document as well as I can remember, ‘can be offered to the press and any other interviewer for discussion’. Yet some people do not remember examples of good RIAs, some others candidly hand-in the pre-packaged list of RIAs, and show me the lines describing why they are good. This happens in a country with some 200 RIAs produced per year. Note that, as I said, the interviews are mostly in the cabinet office and key departments.

The truth is simple: most people talk a lot about RIA and better regulation but do not see impact assessments every day. I am told that they work with ‘the hearts and minds’, ‘cultural change’, and ‘innovations’. I ask about indicators of regulatory quality employed to keep track of a large production of RIA (roughly 200 per year, as mentioned). There is nothing in the public, although in the meantime I have learned that the cabinet office has recently experienced with a performance management tool to create incentives for good practice across departments.

Turning to the Better Regulation Commission, these high-profile personalities do not read RIAs on a regular basis. Instead, they talk to the cabinet office’s specialists, and read the reports of the National Audit Office instead. They are involved in practically every issue of regulatory reform in Britain and Europe, from inspections to the reduction of administrative burdens. Thus, it is not at all surprising that they do not have first-hand information on RIA. Yet their judgment on the overall quality of regulation is quite valued by the core executive, and since RIAs are an important component of better regulation, one would expect more engagement. Finally, the two of us manage to track down the real ‘case handlers’ – those who perform RIAs in the departments, and we start completing the scales and fill in the questionnaires. Even in these cases, however, it is difficult to find examples of RIAs that are particularly good in some specific areas – very few interviewees remembered cases in which a RIA eased problems of
enforcement and implementation, by spotting at an early stage potential problems of compliance and so on.

The UK fieldwork diary suggests a couple of points. RIA is certainly dominant in the discourse around better regulation. The UK is a champion of RIA in Europe, the British are discursive leaders. The government has recently re-designed new written guidance on this tool precisely because it wants more from it. Yet if we move from discourse to praxis the reality is that RIAs are less important that one would think. Interviews with specialists on law-making and pre-legislative scrutiny confirm that the salience of RIA in the policy formulation process varies from low to moderate. Scholars like Ed Page, who have studied policy formulation in Britain from every possible angle, have very few pages (indeed very few lines!) on RIA in their publications (Page, 2003; Page & Jenkins, 2005).

This is perhaps trivial. The interesting point is, arguably, the second: why does the public discourse in the specialized press and in policy communities puts so much emphasis on RIA? Certainly there are lots of people who are daily involved in producing the 200 or so RIAs every year. But why do the others, the ones who do not see RIA, use it as a flagship reference in their public discourse?

This reminds us of the early studies on planning in France. Pierre Muller, Bruno Jobert and their associates went out for their fieldwork with the idea of examining how the French technocrats were planning. Although everyone was enthusiastic about ‘talking planning’, these people – Muller found out - were not engaged in planning. Why on earth where they investing so much time and energy in ‘planning’ meetings and filling in public bodies with ‘planning’ specialists? Muller explained that planning was used to make sense of a new social reality. It was the tool through which a new paradigm or frame of reference (modernization in this case) was diffused – via socialization orchestrated by cultural mediators (Muller, 1995). If this line of reasoning in terms of normative frameworks is plausible (Jobert, 1989), the interesting questions to ask are not the one contained in our interview toolkit. Instead, we should ask questions about the frame of reference that is channeled through RIA discourse. RIA is the carrier (a référentiel?) of broader notions of modernization, risk tolerant de-regulation (Dodds, 2006), and the litmus test that tells us where the political pendulum of regulatory policy is swinging in the UK (Radaelli, 2007).

4. WHO FRAMED THE DUTCH RIAs?

After this somewhat surprising encounter with the British scene, the second country selected for fieldwork is the Netherlands. Like in the case of the UK, there is reason to be confident when we start our fieldwork.

We have taught Dutch civil servants in the past at Exeter University, where we run a course on impact assessment twice a year. Before fieldwork starts,
however, we decide to start looking at a sample of RIAs. The documents of the OECD and the European Commission state that the Netherlands produces RIAs of good quality. Let us read and score some of them, we think.

Language is not a problem. We have a budget for translations. In consequence, we start calling our contacts in the Netherlands and ask for the Dutch RIAs. Time goes by – no RIA shows up in our inbox. We investigate the websites, and cannot go much further than the explanatory memorandum that accompanies proposed legislation. This is fine – we say to our contacts – but where are the RIAs? After all, this is a European champion of better regulation, and its RIAs are universally considered of good quality.

For a couple of weeks, the search for the Dutch RIAs in our department becomes more elusive than the search for the Holy Grail. We know that the Dutch system RIA is organized around a set of checklists that are used flexibly by the teams developing proposals. We realize we have made the mistake of thinking that these checklists produce a document called RIA that is structurally similar to the one we see in the UK.

The reality shows that our conceptualization does not travel well. A ‘proposed legislation desk’ at the Ministry of Economic Affairs in The Hague keeps track of how departments go about using the checklists. The desk asks questions with the idea of helping the department to make progress with their proposals. It hasn’t got the expertise and resources to challenge the calculations made elsewhere. All they want to know at the desk is whether a department has provided a decent answer to the checklist. If it hasn’t, they call and ask for clarification and improvements. The desk cannot stop legislation if it is unhappy about the quality of RIA. In any case the Ministry of Economic Affairs in the Netherlands, is not a major player. The head of the Dutch independent board on the quality of legislation (Actal) was quoted by a newspaper as saying:

“At the moment the Ministry of Economic Affairs is at the bottom of the ‘pecking order’ in The Hague. Actal was terrified at the time that we would be placed under the Ministry of Economic Affairs; that department cannot get anything done”\(^\text{3}\) (our translation)

Going back to our obsession, where are the phantom-like Dutch RIAs? There is no formal document containing the calculations and arguments we usually associate with a RIA. There is no text, although there are numbers, ideas, forms of appraisals that go back and forth from one department to another, from

\(^{3}\) As quoted by *De Telegraaf*, see http://www.overgeld.nl/rubriek/belastingen/1283761/_rsquo;Slimmere_regels_kunnen_miljarden_besparen_rsquo_.html
government to stakeholders, from one desk to another in multiple loops until enough consensus is found. Then the *tourbillion* of ideas, numbers and appraisals settles in the explanatory memoranda. There is no RIA in the Anglo-Saxon, formalistic sense, yet there is a RIA in the sense of appraisal of legislation. How good or bad it is difficult to say. Good for consensus politics in an executive with a skinny majority and electoral volatility, coalition pacts, and effervescent politicians. Bad for us researchers willing to score the Dutch RIAs.

There is another surprising element. The Dutch system is dualistic. One the one hand, there are three checklists on the quality of legislative proposals. As mentioned, they are monitored by a unit in the Ministry of Economic Affairs. In a sense, this element of the system is not unitary, as there are three ramifications.

On the other, the Netherlands has embarked on the most ambitious initiative in Europe to crack down on administrative burdens for enterprises. This is done ex ante and ex post. In the ex-post version, there is a procedure that generates some orders of magnitude of the administrative costs incurred by firms as a result of existing legislation. The procedure is then used to calculate a baseline and deliver simplification plans based on reduction targets. Quality control is performed by an inter-departmental unit called IPAL, situated high up in the hierarchy of the powerful Ministry of Finance. Departments that do not do well in their simplification exercises have to explain to the Minister of Finance why they should deserve more resources when the budget is discussed. The link between simplification and the budget cycle is quite effective.

The ex-ante analysis is performed on proposed legislation. In terms of our definition, it is a RIA. It does not look at benefits, but only at a category of costs. But it is a type of impact assessment. When the burdens enter the equation, control on the quality of the measurement of the burdens is performed by the above-mentioned watchdog, Actal. The latter can and does issue (publicly) negative opinions if the quality is not up to the standards. The civil servants we interviewed for this project feel that Actal is a real watchdog. The result is that in our interviews when people refer to RIA, they can either mean the checklist procedure or the burden procedure. This has led us to add qualifications to all questions. When we ask about whether RIA in the Netherlands has increased the power of the core executive, the respondents give low scores if they think of the checklists, high scores if they have the burdens programme in mind. We try to explain the difference and collect the ‘right’ scores each time. In one case an interviewee decides to provide two scores – one for the checklists and another for the burdens.

In conclusion, the fieldwork in the Netherlands reveals a dualistic system that makes our measures less robust than we thought. We may well consider only the indicators on the checklist, but why would impact assessments on the burdens be excluded from RIA? But if both types are RIAs, we have to qualify our measures of political changes brought about by impact assessment with the
sentence ‘depending on the type of RIA we are considering’. This is still a result, but at the outset we thought we could say something more general on RIA country by country. Instead, the real cases for analysis are the types (of RIA), not the countries.

Another problem with dualistic systems is that they increase the likelihood of comparing apples with oranges. As we cannot collapse the two types into an ‘average Dutch RIA’, how can we compare Dutch and, say, US RIAs? Further, the classic Anglo-Saxon RIA includes consultation as an important mechanism of open governance. It provides data. It enables the stakeholders to articulate their views in the regulatory policy formation process. It makes public administration more ‘intelligent’ by alerting them on how business and citizens may comply and how they will be affected by new regulation. This is fine, but in the Netherlands consultation operates via mechanisms different than RIAs, for example mixed committees in key departments, model firms and so on. However, in our project we want to measure the consequences brought about by RIA by looking at how it provides access to diffuse interests and business interests. But if consultation takes place outside RIA, the respondents will answer these questions by saying that there is no change in terms of this dimension. Yet it would be a mistake to draw conclusions on the quality of consultation in a country with an Anglo-Saxon RIA and a country like the Netherlands. All we can say is that in the former there are changes brought about by impact assessment, in the latter we should not even try to measure how RIA has affected consultation.

Finally, we cannot really perform analysis at the micro-level if the Dutch RIAs do not exist in the form of a written document that provides an explanation of how the problem was defined, what type of consultation was carried out, and how the different options were selected and examined in relation to a set of well-specified decision-making criteria. Our notion of comparability of individual RIA – we have now found out – hinges on an ethnocentric presupposition of what a RIA should be. Obviously, this presupposition is not valid knowledge about what RIAs are!

5. THE DANISH CONUNDRUM

There are several doubts in my mind, but it is high time I went to Copenhagen with the usual batch of questionnaires to be filled. The problem is that most people on the phone do not recognize what we mean by impact assessment: do we want to talk about policy formulation, simplification, reduction of burdens or the impact of the competitiveness agenda on assessment tools? RIA as such is not easy to identify for our interlocutors. Thus, it is already difficult to find a decent sample of people willing to talk because most do not understand what the

---

4 Listening to the interview tapes now, there are exhilarating sessions where I am speaking to two people, their boss gets in a bit late, opens the questionnaire and asks the two colleagues in Danish ‘but what is this RIA that we are supposed to talk about’?
project is all about. A member of my team is fluent in Danish, and after two weeks of efforts he produces a good set of meetings for my trip to Copenhagen.

Here we find a dualistic system like in the Netherlands – one for RIA and one for the burdens. Denmark shares with the Netherlands a phase of political excitement about burdens. The ‘burdens-only’ RIAs are quite elaborated, and the quality control mechanism at the Danish Commerce and Companies Agency is efficient. Indeed, this agency contributes directly to these RIAs by performing some calculations of the size of the burdens for departments and other ministries.

Interestingly, there is written guidance on Danish RIAs. We immediately translate the guidelines, thinking that they capture the essence of the Danish RIAs. This time we are going to find ‘them’ – we giggle. Once translated, the set of guidelines looks similar to Anglo-Saxon written guidance. Accordingly, the first person to be interviewed is the author of the guidelines on Danish impact assessment. After a meeting with him and some other officers in key departments, we find the following. True, written guidance describes a linear process of problem definition, identification of options, analysis of options, checks on specific aspects of proposed regulation, and conclusions. The reality is quite different, however, with informal coordination among departments being far more important than formal-istic RIA exercises. According to one interviewee, ‘if the Danish policy formulation process were to follow the written RIA guidance in its entirety, it would come to a grinding halt in two weeks’.

Other interviewees told me ‘we do not have RIA in Denmark’ – this was even written in their PowerPoint presentation prepared for my visit. Of course, they meant ‘we do not have an Anglo-Saxon formal system of impact assessment, and we do not follow strictly our own OECD-inspired guidelines, we do not read them as a description of how a RIA document should be prepared, step by step…. Hence we do not have RIA’. This is in contrast to the conventional discourse in OECD and EU circles, where the Danish approach to RIA has been presented as good practice for a number of years. If taken seriously, this would mean that we have inserted Denmark in the wrong sample of countries with a long experience of RIA. The reality is even more complicated than that, and stimulates deeper interrogations.

An interesting Danish feature is that, like in the Netherlands, the empirical analysis on which regulatory decisions are based is not really made public. Arguably, one could say that because of this ‘Denmark has no RIA system’. But this feature – according to the interviewees – is supposed to ‘make the system work’? How is it possible? And what is the system? The ‘system’ – we learn gradually – has to be situated in the context of strong ministerial autonomy in the context of coalition executives. Here there is no strong attempt to separate empirical and political analysis. One calculates, consults and changes at the same time. Economic analysis is not pitched against political analysis, neither is
it considered the first step in a rational process that first gathers the evidence and then starts discussing politics. The presupposition that RIA is an empirical exercise that informs a political process is proved wrong in the Danish case. But if this is true, we cannot measure how well empirical analysis supports political choice. Another complication is that this does not mean that policy appraisal is bad or empirically weak in Denmark. Indeed, there is no frustration with this aspect of Danish policy-making, and no evidence that this is the case in the literature we have examined. The quality of appraisal – most interviewees report - is high, although there is no separation of ‘empirical’ and ‘political’ modes. Perhaps this is a case of pragmatism and a plural definition of rationality in politics – by this I mean that rationality can take the form of accurate economic assessment of consequences, but informal coordination and extensive consultation outside RIA are not ‘irrational’. Perhaps there is more complacency than satisfaction – indeed it is hard to imagine that anyone can assess the quality of appraisal if the assessments are not made public. It is difficult to sort out these rival alternative hypotheses (satisfaction versus complacency).

For the time being, we can hypothesize that if the system ‘works’, then we have encountered equifinality: countries can achieve high quality appraisals by pursuing different strategies. What matters is the coherence between institutional structure, administrative styles, and tools. Not the tool alone, but a specification of the tool in a given context. As mentioned, there is empirical analysis supporting decisions, but empirical information and analysis come from multiple sources. There is no assumption that RIA as document and process should be the only (and not even the most important) vehicle for gathering empirical evidence. And no attempt to chain a pragmatic process to RIA formalities either. The observer going out for fieldwork is tricked by the prima facie impression that this is a rather systematic, formal process. But in practice these formal elements are in large part symbolic. People pick and choose from the guidelines, they do not see them as a description of the rule-formulation process.

6. LINNAEUS AND SWEDISH VARIETY

After Denmark and the NL, I thought I had seen enough of dualistic systems and would not be struck by examples of novelty in RIA forms. The fieldwork to Sweden – we reasoned in our team – should not reveal too much incongruence with the previous experience. We had seen enough variety. But in the country where classification was invented by Carl Linnaeus, the forms of Swedish RIAs are incredibly varied and hard to classify.

To begin with, they vary by guidance – it depends if one is doing RIA following the Simplex Ordinance (a checklist on the impact of proposals on small and

---

5 In turn, the impression is originated by the documentation produced by international organizations, self-assessed questionnaires, and tables showing that Denmark has a RIA system with no further qualification.
medium business) or the Agencies Ordinance or the Committee Ordinance. Secondly, they vary by level of governance, that is, committees of inquiry, ministries, and agencies. As we have learned from the literature on this country (OECD, 2007), the Swedish law making process starts with early ideas in the departments about legislation, then there is a bit of collective decision-making in the government about the type of committee of inquiry to establish (parliamentary or bureaucratic). The story goes on: the committee produces a report with something like a RIA in some of the cases (especially if the terms of reference drawn by the government ask the committee RIA-type questions), then there is consultation. Finally, the ministers engage in joint decision-drafting and produce a proposal for legislation with another RIA (this can be the same original RIA produced by the Committee with or without a few added points to justify the choice of a specific option, or a more creative engagement with the RIA exercise). But the agencies also have regulatory power, in which case they produce RIAs according to the Agencies Ordinance. And they also have to follow the Simplex Ordinance.

Thirdly, we have variability on time. The Committees’ RIAs are prepared before consultation, so they are produced at an early stage. As such, they cannot possibly be in the analytical form of, say, cost-benefit analysis. By contrast, when an agency regulates and produces a RIA, there is a pretty specific idea in mind, and economic analysis is more appropriate. The problem is that most RIAs prepared by the agencies are done at a late stage. So either too early or too late, it is never the “right” time for a RIA to appear in the Swedish regulatory policy process!

Fourthly, the Swedish RIAs vary by quality control body and mechanisms. Take the agencies. For the part of the agencies’ RIA involving small and medium business, the quality control unit is Nutek (the agency for economic growth and regional development, which reports to the Ministry for Enterprise), for the other components of the RIA it is the National Financial Management Authority. Turning to the law-making process, the main quality control unit is the better regulation body within the Ministry for Enterprise. At the Ministry of Finance they also have a unit in charge of quality control, looking at the macro-economic and long-term environmental impacts of proposed legislation. Yet all this paraphernalia of quality control bodies does not produce major quality control: there are several controllers but the system is not under control, since no-one can stop a regulation or a bill because the RIA is of low quality. There is no watchdog in the system. The current government is frustrated with the quality of appraisals and, following recommendations made inter alia by the OECD and a small governmental committee of inquiry, is changing the systems.

Fifthly, the RIA system varies by resources. Take quality control: the Financial Management Authority performs quality control with two people, not even full-time, Nutek has some 12 people. So we have 12 people looking at one minor aspect of the RIAs (the impact on small business) and less than 2 FT-equivalent
that are in charge of all the other quality aspects of the Swedish RIAs produced by the agencies. Agencies do not have enough economists to perform proper RIAs; at KEMI, the agency of chemical regulation and inspections, there are 200 people, of which some 3-4 are economists. Time is another important resource for RIA, and in this case too there is variability. Sometimes the agencies do not have the time to perform proper analytical RIAs, but the committees may have three years – during which however they produce a policy investigation, not a RIA.

The sixth dimension of variation concerns the formal-informal divide. Here we are back into familiar Dutch-Danish territory. The appraisal of proposals is quite solid, but it does not hinge eminently on analytical RIAs. There are several informal patterns of interaction, forms of analyses, consultation etc that enable the policy makers to appraise the quality of proposed regulation. These elements are more important than RIAs. In the agencies, economic studies are quite lengthy, and are used in formulating proposals. The RIA is instead a short document in which the questions posed by the checklists are answered to. So, is RIA bad in this case because all the meaty analytical work falls in a 200-page report and consultation falls outside the RIA? Or is the 200-page document the RIA and the questions of the checklist a bureaucratic add-on where someone ticks the box? Should we be concerned about the low quality of RIA (at this point we do not even know what is the “RIA” we are talking about….) or about the low quality of policy appraisal?

The point about committees made earlier on brings us to the final remark about variability. There is also variability in terms of policy-making styles. Given their nature, the preparatory work done in the committees is quite political, based on multiple feedback loops between problem definition, analysis of options, exploring the boundaries of what is politically feasible, performing some tests, going back to the drawing board. The chair of the committee has an important role to play not only as agenda-setter, but also at crucial stages in these feedback loops. He or she can moderate the debate, but also suggest a solution. In this political magma, clumsy solutions (in the sense of Verweij and Thomson) are efficient: there is no single metric that can produce a solution, yet the intelligent use of conflict brings democratically acceptable results (Verweij & Thomson, 2006). To superimpose to the magma a single-metric, linear tool like RIA is a violation of the fundamental aspects of committee governance in the Swedish context. Clumsy solutions need different metrics in order to use conflict intelligently.

This contradiction between RIA and the process through which solutions are discovered may be much less important if the committee is staffed by a single civil servant, or in the case of agencies. Here the style is more technocratic and regulatory analysis sits more comfortably with the logic of a RIA – at least this is my hypothesis, and it would be interesting to test it on a sample of RIAs across committees and agencies.
The first stage of fieldwork is finished. I forgot to compare how the governments feel about their RIA systems. Let us re-cap then. In the UK, there is frustration with the quality of impact assessment - the government changed written guidance and format for RIAs in 2007. In the Netherlands, they are very happy with the burdens-only RIAs, but unhappy about the performance of the standard RIAs monitored by Economic Affairs. In Denmark, the key departments and DCC are quite happy with their systems. In Sweden, as mentioned, the government is convinced that different set of guidelines and the poor quality control have not done a very good service to the cause of policy appraisal.

7. TO CONCLUDE: **RECULER POUR MIEUX SAUTER?**

As mentioned in Section 1, the lessons can be grouped in three categories. The first lesson is about the relationship between conceptual analysis and measurement. Having established that the former comes before the latter (Brady & Collier, 2004; Sartori, 1970) what is the next step? What is the best way to proceed in terms of concept formation? Here are some examples of the questions we face. Is a Dutch RIA confined to the explanatory memorandum accompanying legislation – this being the only public component of the process leading to legislative proposals? If, alternatively, a Dutch RIA is more than this, but the material outside the explanatory memorandum is not in the public domain and is not organized according to a template recognized by us researchers, can we include in a concept something that we cannot see empirically (unless we have access to some PCs in The Hague)?

This observation brings us to an important distinction, namely the difference between species and genus. We do not know (as yet) how to classify *per genus et differentiam* in this field of research. A genus is a taxonomic group containing one or more related species. Animals are often classified using a two-part name (in Latin): the first part is the genus, the second identifies the species. So for example we have Canis-Lupus, Canis-Latrans and so on. Beyond the genus, we have the family, made up of different genera. By following this simple set of rules, we can compare pretty safely – otherwise there is a risk of miscomparing (Sartori, 1991).

Turning to better regulation and RIA, current research has not established what the genus is and what the species are. One sensible way to proceed is look at the different species. We need to create a typology of RIA, identifying the common set of properties that apply to each species. We may have RIA-benefit cost analysis, RIA-administrative burdens only, RIA-techno-political appraisal (the Swedish committees with parliamentary participation), but is ‘RIA-informal coordination and appraisal’ yet another species of RIA or does this belong to the genus ‘administrative procedures used at the pre-legislative scrutiny stage’?
Assuming we manage to establish RIA types, is type 1 common to a specific family of nations, or can we have type 1 and type 2 within the same country? Shall we compare by type or by country? What is the genus that contains the various RIA species? How does it differ from others – I mean what is the genus that contains the non-RIAs? Note that different genera constitute a family. Future research will have to ‘go up’ and identify the family of the genus RIA, arguably this would be the better regulation family.

Finally, there are issues of multi-level analysis. Typically, we handle RIA as if we were handling units of one type (hence we sample RIAs as if we were sampling individuals, think of a researcher analyzing pupils). However, if we start using RIA types, we have to account for complex data sets (with units classified by type of RIA, type of department, countries, and perhaps even more units). To go back to the example of individuals, this would be the same as shifting from pupils to the measurement of educational achievement, in which the units are pupils, teachers, and schools. The conceptual and statistical properties of multi-level analysis are different from the analysis of observations of units of analysis of one type (Snijders & Bosker, 1999).

The second category of lessons is about the asymmetries between demand and supply of research, or the relationship between the community of researchers and the policy-making world. The strong demand coming from international organizations and governments has led researchers to move too swiftly and perhaps not in the right direction. The clients want more research on RIA, thus we badly want the RIAs to be there. We design RIA indicators and sophisticated templates for data collection such as DIADEM but, as our discussion within the ENBR High Level group of experts on methodology has shown, we are not sure whether there are all these RIAs ready to be scored. We may have more stuff in the non-RIA genera than in the RIA genus. We may have more unpublished and informal assessments of policy than standardized RIA documentation.

One way to balance the relationship between demand and supply is to engage in problem re-definition. The demand of the client has to be reformulated, and researchers may be more helpful if they can usefully contribute to the re-calibration of the key questions. To illustrate, the client may want to know more about comparative statistics on RIA. This is a very sensible policy question. But at this stage, we may need to transform this into a research question, and go back to the client with questions on the emergence of different RIA types. Before we score and provide measures, we need to explain how different countries adopt the same language of better regulation and RIA but implement very different types (Radaelli, 2005). Key research questions are: what is the constellation of actors, problems, solutions that leads to type n? Is this type efficient or dysfunctional to country X, given institutional context, administrative procedures, and pressure group politics in X? The question about emergence of different types may be handled by Boolean algebra, looking at whether outcome
‘type A’ is associated to conditions X and Y or conditions X and Z. One way to explore these conditions is qualitative comparative analysis (Ragin, 1989), a tool that has not been used as yet in better regulation research.

Questions that obviously come to mind are essentially about the comparative statics and dynamics of the adoption and evolution of better regulation policy instruments. For comparative statics, we need to explore how a shock alter the overall equilibrium on which better regulation is based. Think of the introduction of an innovation in the tool-kit of better regulation like the standard cost model. Comparative statics can tell us if regulatory innovations such as the standard cost model change the overall balance between and supply of better regulation – or the political architecture of who gets what. Turning to dynamics, the following issues need to be investigated. How do different types of RIA co-exist and evolve, with species becoming gradually dominant – for example we may see cost-benefit RIAs being overcome by administrative burdens assessments? And then the ‘why’ question: if they evolve, there may be a causal explanation of the changes that we must try to unveil with a theoretical model and then control empirically.

The relationship with the client can also be rebalanced by providing more useful templates for policy suggestions. At the moment, the new EU Member States that are experimenting with the introduction of RIA look at the OECD Anglo-Saxon template for inspiration. And they find it difficult to implement. One option (more important for researchers than for policy makers) is to consider the reasons why countries like the Netherlands and Denmark have not implemented an Anglo-Saxon type of RIA and see if these reasons apply to the new members of the club as well. Another option is to think about context-based policy recommendations that recognize explicitly that the ugly duckling that does not conform to the usual template may be a beautiful swan. Perhaps some countries do not have to important the classic template, perhaps they have to look at countries like Denmark or the Netherlands for inspiration, if their political-institutional context about consultation and policy appraisal is the same.

The whole business of adaptation, in turn, is tricky. On the one hand, implementation is always a matter of transforming and editing policy templates in the light of local political and administrative conditions. But we do not know under which conditions ‘adaptation’ means that the RIA smoke does not produce any fire at all, because it has been ‘adapted’ so much that it is no longer challenging the status quo, and henceforth does not improve on the current state of affairs.

A final remark on the relationship with the policy-makers - this time from the point of view of how researchers approach their interviewees. Typically, interviewers approach policy makers by accepting their definition of what is a RIA and relying on the self-assessed questionnaires that governments send to the OECD and the European Commission. In consequence, we take for granted that there are RIAs in, say, Portugal, if this government has answered YES to previous OECD
questionnaire. All we want to establish is to have access to the RIA – which often triggers a long discussion on the written guidance in Portuguese, complemented by a loose set of data and information that according to the interviewee is the RIA system (after all, this is why the answered YES to the questionnaire). Perhaps it would be better to challenge the interviewees directly with clear definitions of what we expect to be a RIA in our research project, and, if this is the case, make them realize that they do not have RIAs, or do not use cost-benefit analysis, or have a practice of consultation that is exogenous to better regulation. By doing so, we can increase awareness and contribute to a common understanding across countries. This would be a much more reflective way to conduct interviews and carry out fieldwork.

Partly related to the other two lessons, the third lesson is about who defines our object of study. We – the researchers – did not invent better regulation. The policy-makers did. We have seen several cases in which the policy-makers produce policy innovations, and literally invent new ways to go about the governance architecture of complex societies – the open method of coordination is yet another example. Researchers are first taken aback, then after some months or years start explaining to the policy-makers what they (the people in the ‘real world’) have done. By doing so, the researchers make sense of reality with the same vocabulary used by the practitioners.

Now, there is no doubt that better regulation is an excellent term for political action. It mobilizes politicians and bureaucrats. It resonates with ‘improvement’ and notions like the modernization of the public sector. It brings with it a discourse that legitimizes some changes. It silences those who are opposed to innovation – it is easy to make an argument against, say, tax harmonization, but who can be against better regulation and make an argument for .... ‘worse’ regulation? However, the researchers are not at ease with terms that are normatively loaded. It is difficult to measure quality in better regulation, because this would lead us to talk about ‘good better regulation’ and ‘bad better regulation’. The lesson, therefore, is that before we start churning out data, we need to think carefully about our scientific vocabulary. This is why in recent work I prefer to use the notion of meta-regulation instead of better regulation (Radaelli, 2007).

Not only does better research need its own vocabulary. It also needs to choose the dependent variable(s) on the basis of solid theories and models that enable us to predict causality. The sad state of affairs in RIA research is the disproportionate attention paid to scoring cards and the neglect of theory, including, paradoxically, theories of regulation. The problem is compounded by the tendency to look at RIA in isolation from the other components of the policy formulation process, most importantly perhaps, the various ways in which administrative procedure is defined across Europe. A more solid theorizing and an integrated analysis of RIA would put us on more solid ground. We would be able to go out for fieldwork with a predictive causal model in mind (no matter how
rudimentary it can be at this stage), instead of puzzling over vocabularies we do not control, RIA observations that mean different things in different contexts, and the impossible task of comparing apples and oranges.

This is perhaps a sad way to conclude a paper for a Network dedicated to the production of the first pan-European dataset on impact assessment, but at this stage a pause for reflection, I am convinced, will help us to reculer pour mieux sauter.

References


