Moving boundaries in transboundary air pollution co-production of science and policy under the convention on long range transboundary air pollution

Willemijn Tuinstra\textsuperscript{a,b,*}, Leen Hordijk\textsuperscript{c}, Carolien Kroeze\textsuperscript{a}

\textsuperscript{a}Wageningen UR, Environmental Sciences, Environmental Systems Analysis Group, Wageningen, The Netherlands
\textsuperscript{b}Netherlands Environmental Assessment Agency, European Air Quality & Sustainability, Postbus 303, 3720 AH Bilthoven, The Netherlands
\textsuperscript{c}International Institute for Applied Systems Analysis, Laxenburg, Austria

Received 26 July 2005; received in revised form 1 March 2006; accepted 6 March 2006

Abstract

This article focuses on the science–policy interaction in international negotiations in the context of the United Nations Economic Commission for Europe's Convention for Long Range Transboundary Air Pollution (CLRTAP). It addresses the question how participants in the assessment process divide and co-ordinate work between science and policy and how this enhances credibility, legitimacy and relevance with multiple audiences. For this purpose the article combines an analytical framework to approach effectiveness of scientific assessment in policy making, with the notion of boundary work and co-production of science and policy. The article argues that knowledge produced within the CLRTAP process and the institutional setting in which this knowledge production takes place cannot be separated from each other. Furthermore credibility, legitimacy and relevance are to a large extent determined by boundary work in an early stage of the process. At the same time boundary work has to take place continuously in order keep the assessment process credible, legitimate and relevant for new audiences. The application of a combined framework for analysing credibility, legitimacy and relevance and for analysing boundary work turns out to be helpful in describing in detail what happens in practice at the science–policy interface. In particular it helps to address the question of the way participants in the assessment process divide and co-ordinate work, how this shapes design elements and how this enhances credibility legitimacy and relevance of an assessment.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: LRTAP Convention; Science-policy interface; Boundary work

1. Introduction

In the field of policy making for air quality in Europe a long tradition exists in using scientific information to support negotiations and decisions (e.g. Hordijk, 1991; Gough et al., 1998; Tuinstra et al., 1999; Castells and Ravetz, 2001; Eckley, 2002; Sundqvist et al., 2002; Lidskog and Sundqvist, 2004). This article focuses on assessment processes in air quality policy making in Europe. Assessment processes are intensive communication processes between scientists and policy makers, that aim at integrating knowledge from different scientific disciplines e.g. economy, soil-science, ecology, meteorology and other knowledge sources in such a way that it provides useful information for decision making (e.g. Farrell et al., 2001). Assessment processes are embedded in a variety of institutional settings, within which scientists, decision makers and other stakeholders communicate to define relevant questions for analysis, mobilise certain kinds of experts and expertise, and interpret findings in particular ways (Farrell et al., 2001).

Various notions of the relationship between science and policy are of significance for current questions and debates
with regard to the role of science in public policy. What demands have to be made on the communication process between scientists and policy makers in order to create assessments that indeed provide useful information to decision making? When are assessments considered to be useful? How can scientists maintain their credibility when they engage in policy issues? These are actual questions which are of concern at various levels of the science and policy domain. At the same time social science research has shown that a one-directional linear relationship between science and policy in which science provides objective “answers” for policy is an illusion (see e.g. Jasanoff and Wynne, 1998). What counts as an “answer” depends on how the problem is framed in the first place. Scientific knowledge is not independent from cultural factors, society and policy. Furthermore, increased scientific knowledge raises new questions or can make decisions even more difficult (e.g. Rayner, 2006). In this particular case we focus on how participants in an assessment process within the context of air quality policy making in Europe divide and co-ordinate work between science and policy and how this enhances credibility, legitimacy and relevance of assessments with multiple audiences. The article addresses the role of science–policy interactions in shaping assessment frameworks and assessment processes.

The analysis in this article builds further upon the concept of effectiveness as developed in the Global Environmental Assessment (GEA) project (Farrell et al., 2001; Farrell and Jäger, 2005; Mitchell et al., in press). This concept considers effectiveness as an emerging property based on three qualities that participants and users attribute to an assessment: credibility, legitimacy and relevance. These qualities are co-determined by the characteristics of the assessment itself, the characteristics of the users of the assessment and the context in which the assessment takes place. In this article we refer further to this framework for considering effectiveness as “the GEA-framework”. We adapt the GEA framework and take it further, connecting it to the concept of “boundary work” between science and policy.

Our starting point is that it is not easy to draw a sharp line between scientific and policy making activities in an assessment process. Neither can scientists’ or policy makers’ roles as actors in such processes always be precisely defined. Negotiation takes place about the identity of practices (e.g. “science” and “policy”) and actors (e.g. “scientists” and “policy makers”) and their collaboration. This practice of maintaining and withdrawing boundaries between science and policy, shaping and reshaping the science-policy interface has been referred to as “boundary work” (cf. Jasanoff, 1990; Gieryn, 1995; Halfman, 2003).

In this article we combine the features of the GEA framework with the notion of boundary work into a framework for analysis which adds to the literature because it has a special focus on the science–policy interface. This helps us to improve our understanding of co-production of science and policy: the simultaneous development of the problem framing, division of labour between science and policy, the assessment framework and the policy framework.

The empirical data for this article come from the science–policy interplay in the negotiations within the framework of the United Nations Economic Commission for Europe’s (UN-ECE) Convention for Long Range Transboundary Air Pollution (CLRTAP). CLRTAP is a well-known case for scholars exploring the role of science in policy making (see e.g. Gough et al., 1998; Grünfeld, 1999; Tuinstra et al., 1999; Bäckstrand, 2001; Castells and Ravetz, 2001; Eckley, 2002; Sundqvist et al., 2002; Farrell and Keating, 2005; VanDeveer, 2005). It is an inviting case to study boundary work and co-production. Not only because scientific information has been playing an important role in the negotiations, but also because policy-makers and scientists have been dividing work in various ways on various occasions and also change roles regularly (see e.g. Eckley Selin, 2005). This offers anchors for lessons for other policy areas. The analysis in this article is based on (1) the study of official documents and reports of meetings of various bodies operating in the science–policy interface within the Convention as well as informal documents, (2) interviews with delegates, chairmen of working groups, and scientists involved in the process, as well as interviews with scientists and civil servants who were not directly involved but followed the process from a distance, and (3) observations in official meetings of CLRTAP-bodies as well as observations of working processes and informal meetings of involved scientific institutes.

The next section presents our framework for analysis. Section 3 presents the history of the interaction between science and policy within CLRTAP. Section 4 applies the framework to explore the case. The final section draws conclusions and points to wider implications of the study for other policy areas.

2. The role of boundary work in shaping assessment processes: A framework for analysis

2.1. Introduction

In this section we describe how we combine the framework for analysis of effectiveness as defined in the GEA-project with the theoretical concepts of boundary work and co-production in order to help us focus on the science–policy interface. In the original GEA framework the science–policy interface is one of the design elements of
an assessment. However, our starting point is that the design of the science–policy interface determines other design elements of assessments like e.g participation, treatment of uncertainty, and treatment of dissent. The ways by which participation, treatment of uncertainty and treatment of dissent are organised are inherently part of the science–policy interface. We are interested in what happens at the science–policy interface and how this shapes the design of the assessment. The concept of boundary work helps us to examine the science–policy interface further. Therefore, for the purpose of this article, we will not handle the science–policy interface as being one of the design elements of an assessment (as in the original GEA-framework), but we approach it as the site where boundary work and co-production of science and policy take place. Section 2.2 introduces the original GEA framework in more detail and Section 2.3 introduces the notion of boundary work.

2.2. The GEA framework

The GEA framework (Fig. 1) starts from the view that assessments are first of all communication processes (Farrell and Jäger, 2005; Clark et al., in press). The framework acknowledges that it is not really possible to define effectiveness, because effectiveness relates to the achievement of goals, and goals of various participants in the assessment differ among each other. No single, objective measure of effectiveness has been or can be established. This means that many ways to assess effectiveness exist, rather than none. The GEA framework focuses on impacts of assessments. Assessments that are considered to have impact also have three important qualities attributed to them by various audiences: (scientific) credibility, (political) legitimacy and (policy) relevance (Clark et al., in press). Credibility refers to the scientific and technical believability of the assessment to a defined user of that assessment. Legitimacy refers to the political acceptability or perceived fairness to a user of that assessment. Relevance refers to the extent to which an assessment and its results address the particular concerns of the user (Farrell and Jäger, 2005). It is important to note that no straightforward way exists to ensure relevance, credibility and legitimacy of an assessment. These assessment qualities are partly dependent on each other and there are complementarities and trade offs between them. Furthermore, relevance, credibility and legitimacy, are viewed differently by different actors, and therefore it will not be possible to design an assessment process in such a way that it will be relevant, credible and legitimate for all actors in the same way. Assessments are having impact only if they are sufficiently relevant, credible and legitimate according to multiple audiences simultaneously (Cash et al., 2002).

The credibility, legitimacy and relevance of an assessment are not only determined by the characteristics of the assessment itself. Eckley (2001) distinguishes two other ultimate determinants next to the assessment characteristics: the historical context of the assessment and the characteristics of the users of the assessment. The context of an assessment includes the position of the issue on the policy agenda and the characteristics of the issue domain itself. User characteristics are, for example, their interest in the issue, resources to engage in the assessment or to use the results, and the openness to different sources of advice (Eckley, 2001).

Relevance, credibility and legitimacy with multiple users, can be enhanced if context and user characteristics are taken into account in the design of the assessment. Assessment characteristics are the practical result of the design, taking into account the context and user characteristics. The effectiveness of an assessment is thus a function of the interaction between assessment characteristics and the social and political context within which the assessment is conducted (Mitchell et al., 1998).

Farrell and Jäger (2005) identify various design elements, of which they view as the most important in assessing the effectiveness of assessments: (1) initiation and goal,
(2) science–policy interface, (3) participation, (4) treatment of uncertainty, and (5) treatment of dissent.

The *initiation and goal* of an assessment are important because the assessment is to a considerable extent shaped by the origin and initial goals of the assessments: by whom and why the assessment was set up and what the organisational context of the assessment is. The initiation and goal of the assessment will also to a great extent determine the framing and the focus of the assessment. The *science–policy interface* refers to the organisation of the interaction between science and policy. *Participation* refers to individuals and organisations involved in the process and how and when they participate. *Treatment of uncertainty* and *Treatment of dissent* refer to how uncertainty is being managed and how different and opposing insights are being dealt with. Other design elements mentioned in GEA publications are scale, framing, capacity, quality control and transparency (Farrell and Jäger, 2005).

### 2.3. Boundary work at the science–policy interface

As pointed out above, our starting point is that the way various design elements are organised in an assessment is inherently part of the science–policy interface. Therefore the science–policy interface is central in our framework (see Fig. 2). The concept of boundary work helps us to further examine the science–policy interface.

An important aspect of science–policy communication in assessment processes is the negotiation of the division of labour between science and policy. Negotiation takes place about the identity of practices (e.g. “science” and “policy”) and actors (e.g. “scientists” and “policy makers”) and their collaboration. This process of maintaining and redrawing boundaries between science and policy, shaping and reshaping the science–policy interface has been referred to as “boundary work” (cf. Jasanoff, 1990; Gieryn, 1995; Guston, 2001; Halfmann, 2003). Note that this fluid image of the dynamics of the science–policy interface is different from an image in which a “gap” between science and policy exists that has to be bridged, or a manifest boundary between science and policy that has to be crossed. The negotiation and establishment of the boundary itself and the definition of science and policy is part of the science–policy communication process. It is through boundary work that boundaries are made “real”.

Boundary work has two sides: a *demarcation* side separating two actions or groups by defining distinguishing characteristics and prescribing proper ways of behaviour for e.g. science and policy, and a *co-ordination* side defining how the two relate to each other by defining proper mutual conditions of exchange (Halfman, 2003). Boundary work leads to a division of labour between science and policy.

Halfman (2003) provides a vocabulary to describe this division of labour which is useful for our study. Halfman distinguishes three forms in which the division of labour can be embodied: *texts, objects and people*. *Boundary texts* (or *language or discourse*) refer to the way actors distinguish between science and policy in spoken and written text and define respective roles. *Boundary objects* refer to the tools that actors use, e.g. computer models, concepts or measuring standards, for knowledge production in a policy setting. *Boundary people* refer to networks of “scientists” and “policy” makers that are formed or individual people who through their position or actions mark a boundary between science and policy. Together, texts, objects and people form the boundary configuration between science and policy which is constructed throughout various stages of the communication process between science and policy within the context of a particular issue domain.

For the purpose of our study, we are not only interested in the division of labour or boundaries between science and
policy and how they are negotiated. We are also interested in the knowledge eventually produced once the boundaries are drawn, or, for that matter, in the changes in knowledge needs and knowledge production when the boundaries change again. We are therefore interested in the question how boundary work shapes assessment processes. In the science–policy interface, through boundary work, knowledge is produced, and simultaneously the social structures to produce this knowledge are being organised and the scene is being set for the framing of the policy problem and the organization of dealing with the problem. Within the field of science studies, the term co-production is used to refer to processes that connect the production of knowledge with the organization of policy-making (Shackley and Wynne, 1995; Jasanoff, 1996; Miller, 2001; Jasanoff, 2004). Boundary work, the division of labour between science and policy, is part of this process.

2.4. Relating boundary work to the design of assessments in order to study co-production

For our analysis of the UN-ECE CLRTAP assessment processes we combine the framework for analysis of effectiveness as defined in the GEA-project with the theoretical concepts of boundary work and co-production. We will describe boundary work in terms of texts, objects and people. We will address the question how boundary work did enhance credibility, legitimacy and relevance of assessments in the CLRTAP process. More precisely: how did participants in the assessment process divide and coordinate work between science and policy; how did this shape design elements (initiation and goal, participation, treatment of uncertainty, treatment of dissent) of the assessment and how did this enhance credibility, legitimacy and relevance with multiple audiences (see Fig. 2). By addressing these questions we aim to improve our understanding of co-production within CLRTAP: the simultaneous development of the problem framing, division of labour between science and policy, the assessment framework and the policy framework.

3. A brief history of CLRTAP: from flat-rate and single compounds to multi-pollutant multi-effect

This section gives a short overview of the history of science–policy interaction in the development of protocols and assessment frameworks in the context of CLRTAP. More detailed accounts of the history of CLRTAP and the environmental concerns, negotiation processes and scientific programmes that preceded it are given by e.g. Tuinstra et al. (1999), Grünfeld (1999), Bäckstrand (2001), Wettestad (2002), VanDeveer (2004), Menz and Seip (2004) and Sliggers and Kakebeke (2004).

The history of CLRTAP goes back to the 1972 Stockholm Declaration of the United Nations Conference on the Human Environment which states that nations have “the responsibility to ensure that activities within their jurisdic-

2In 1998 protocols on heavy metals and on persistent organic pollutants were signed as well in the framework of CLRTAP, but these protocols will not be discussed in this article.
quality-controlled information about the country-specific situations are prepared. The outcome of this scientific assessment process is submitted to the Working Group on Strategies, whose members are civil servants, representing governments of parties to the Convention, which uses this information to assist its negotiations on further emission control agreements.

3.2. Towards the use of “critical loads” and an “effect-based” approach and the principle of cost-effectiveness

Towards the end of the first round of protocols, which imposed uniform reduction requirements to all Parties, the so-called “critical loads” approach emerged as an overarching concept for the scientific assessment. “Critical loads” have been defined as the maximum exposure to one or several pollutants, at which according to current knowledge no harmful effects occur to sensitive ecosystems in the long run (Nilsson and Grennfelt, 1988). With fully quantified ecological damage functions for the relevant ecosystems not being available, critical loads were simplifications representing a steady-state “no-damage” threshold. This simplification turned out to make the information collected by the Working Group on Effects operational for practical strategy development. Operationalisation of the concept of critical loads for practical strategy development was further facilitated by the acceptance of the “cost-effectiveness” principle, used to identify options for emission reduction leading to an efficient reduction in critical load exceedance. This combined approach developed into the major concept for determining the environmental ambition level and negotiating the resulting emission reduction requirements for the Oslo Protocol, in 1994. Several authors have pointed to the importance of the development of the critical loads approach as a common basis for air pollution control both within the policy and science communities. Patt (1999) and Bäckstrand (2001, pp. 125–144) respectively give a detailed description and discourse analysis of the emergence of the concept. Bäckstrand notes that its success is attributed to its capacity “to create a meaningful framework and to invoke common sense notions of adequate approaches to environmental policy-making” (Bäckstrand, 2001, p. 136). According to Grennfelt and Hov (2005), it was a driving force for scientific research and policy development, and provided a “common concept” for science and policy (Grennfelt and Hov, 2005, p. 4). Also Sundqvist et al. (2002) note that the critical loads concept has served as an important tool for connecting scientific knowledge to policy-making. They show that the concept has different meanings for the involved actors, which include heterogeneous views on the boundary between science and policy.

3.3. A joint effort for data collection

The application of the critical loads concept and of the cost-effectiveness principle went together with a major coordinated effort of data collection by the parties to the Convention in the fields of energy, emissions, technology, control costs, atmospheric dispersion and mapping of critical loads.

The collection and mapping of data on critical loads was and still is coordinated by the Coordinating Center for Effects (CCE) which was established at the National Institute for Public Health and Environment (RIVM) in the Netherlands. In all parties to the Convention national focal centres were established to provide this data. A unified methodology was developed in order to ensure international harmonization. A mapping manual, workshops with specialists from all over Europe, and training sessions contributed to the harmonization. The mapping work was an iterative process and the maps were regularly updated (Hettelingh et al., 1995).

Atmospheric dispersion was assessed by models developed by EMEP with its meteorological synthesising centres and chemical coordinating centres in Oslo and Moscow. The models of EMEP relate deposition in grid-cells, which are squares laid over a map of Europe, to emissions in each European country. The EMEP models employ official emission inventories provided by national governments and evaluated by the Task Force on Emission Inventories, use actual meteorological data and are being continuously verified against measurements (Lövblad et al., 2004).

3.4. The use of integrated assessment models

To combine and analyse this complex information in a consistent and efficient way, computer simulation models (Integrated Assessment Models (IAMs)) were used by the Task Force on Integrated Assessment Modelling (TFIAM) of CLRTAP. This Task Force was established in 1986 before the actual negotiations on the Oslo Protocol began. Its task was to “to explore the possibilities to develop an analytical framework for a regional cost-benefit and cost-effectiveness analysis of concerted policies to control air pollution” (UN-ECE, 1986). At that moment several models at various institutes in Europe were under development.

TFIAM discussed and compared various model approaches and made suggestions for further development. For the actual preparations of the negotiations three different models, ASAM (Imperial College, London), CASM (Stockholm Environment Institute (SEI), York) and RAINS (International Institute for Applied Systems Analysis, Laxenburg, Austria) were used in the Task Force to calculate different scenarios with different policy options, which were presented to the Working Group on Strategies. As far as possible the models were run with the same data, provided by the parties to the Convention. The output of the model runs took the form of levels of emission reductions per country and the resulting percentages of ecosystems protected. The final calculations that served as a starting point for the negotiations of the second sulphur protocol, the Oslo protocol (1994) were done by
3.5. A success story in science–policy interaction?

Generally the science–policy interaction which took place within CLRTAP has been described as a success story both by policy makers, scientists and analysts. As factors for this success are mentioned: the direct link between relevant science and policy preparation; multiple scenarios that present policy options; accessibility of science to all participants; cost effectiveness analysis rather than cost benefit analysis; strong personal networks, a science–policy network with a strong memory (Hordijk, 1991; Tuinstra et al., 1999; Maas et al., 2004). Others describe the important role of the framework offered for countries to provide their own data, the consensus based way of working (Gough et al., 1998) and the ability to describe the important role of the framework offered for all participants enhanced credibility, legitimacy and relevance with multiple audiences.

4. Exploring boundary work within CLRTAP

4.1. Introduction

We analyse the role of boundary work (becoming visible in texts, objects and people) in shaping each design element as identified in the GEA framework (initiation and goal, participation, treatment of uncertainty, treatment of dissent) (see Table 1). Each sub-section addresses one of the design elements of the assessment process and by presenting examples analyses in what way credibility, legitimacy and relevance have been enhanced in the assessment process. The examples presented are meant to be illustrative and not to be exhaustive. The analysis focuses mostly on the period between the early 1980s and early 1990s with some excursions to developments in recent years.

4.2. Initiation and goal

4.2.1. Early boundary work at the level of international diplomacy

If we look at the beginnings of CLRTAP and the years preceding its negotiations we see that co-operation between scientists and policy makers started slowly. The scientific community identified the issue of air pollution first, formulated the first framing of the problem and together with environmental NGOs had put the issue on the agenda. See also e.g. Björkbom (1999) and Ågren and White (2004). The 1975 Helsinki conference had framed the air pollution issue as a policy problem for international co-operation (CSCE, 1975). The scientific work of Norwegian and OECD programmes eventually developed the issue further which paved the way for the establishment of EMEP and subsequently CLRTAP. See also e.g. Grünfeld (1999), Bäckstrand (2001) and VanDeveer (2005).

3See for a detailed account of the various scenarios calculated, the models used and the development of approaches, Tuinstra et al. (1999), Gough et al. (1998), Castells and Ravetz (2001) and Maas et al. (2004).
4.2.1.1. Texts. Actors involved in the early days, both those who represent scientific institutes and those who were involved at the ministries indicate that in the early days of EMEP civil servants were not really interested in EMEP: why would scientific information be needed for policy making? Thus in the early days those civil servants made a clear demarcation between what in their view belonged to their responsibility and what to the responsibility of scientists. Only after the Stockholm Ministerial Conference on Acidification of the Environment (1982) civil servants became more open: political legitimacy paved the way for policy relevance of EMEP work. See also Kakebeeke et al. (2004).

4.2.1.2. Objects. The importance of role of scientific data evolved slowly and step by step: the first step was to make sure that an operational Europe-wide (both East and West) monitoring network for measuring air pollutants was available. Estimation of emissions was a next step. Tables were used showing the extent to which regions were sources or receptors of pollutants. Those came from so-called source-receptor matrices produced and used by the EMEP models. The matrices in fact show which countries are responsible for pollution in certain other countries. The use of those “blame-matrices” was only received very hesitantly in the policy community (Schneider and Schneider, 2004). The availability of those matrices would limit space for manoeuvring in the negotiations. Policy makers felt that the existing division of labour between science and policy was being contested. The blame-matrices were therefore carefully introduced by the EMEP steering body at first only to enable a general idea of the air quality situation in Europe. Slowly EMEP proceeded in building credibility and legitimacy in a continuous interaction with policy-makers, negotiating and establishing areas of responsibility for the policy domain and the science domain and establishing the position of the source-receptor matrices.

4.2.1.3. People. After the adoption of CLRTAP in 1979 it took quite a long time to negotiate the 1983 protocol for EMEP. The context of the Cold War period made that the exchange of data on e.g. emissions of pollutants or activity data of industry was a sensitive political issue. Not only the parties had to build up trust in each other’s data, also ways had to be found to present and share information which could be strategically sensitive, like the location or specifications of power plants (Nordberg et al., 2004). The setting of the discussion on “scientific” issues initially was highly “political”. The Steering Body of EMEP itself at the beginning was established as a political body like other bodies under the UN-ECE. This meant that permanent representatives of member countries of the UN-ECE were present at the EMEP Steering Body meetings in Geneva, though mostly scientific and organisational issues were being discussed. However it was decided after a few meetings that the Steering Body was a scientific body, in which no permanent national policy representatives were needed (Kakebeeke et al., 2004). That was a clear action of demarcation and co-ordination between science...
and policy. It implied the redefinition of the identity of an official body with consequences for who belong to that group and who do not. This proceeded in quite a smooth way as it followed from the actual content of the discussion and work of the group, as such enhancing the legitimacy of the work of the group.

4.2.2. Early boundary work at the level of model development

An eye-catching feature in the science–policy interface in CLRTAP is the prominent role of the RAINS-model which was one of the integrated assessment models used in the process. It is interesting to have a look at the way members of the RAINS team took efforts to enhance the relevance of the model for policy makers and how policy makers reacted to this.

4.2.2.1. Texts. The UN-ECE seemed to be the appropriate forum for the RAINS model to be used, because it involved countries from Eastern and Western Europe and the host institution of RAINS, IIASA, was a product of East–West collaboration itself. This East–West background of IIASA has been important for the legitimacy of the use of the model, as it was not representing the view of one country (Hordijk, 1991; Farrell and Keating, 2005).

In 1983 the RAINS team paid a first visit to the Air Pollution Unit of the UN-ECE to present the possibilities of the RAINS-approach. However, initially UN-ECE was not very enthusiastic, because they could not quite see what it would add to the EMEP models and because IIASA was neither party to the convention nor a recognised NGO in air pollution issues. However, some members of the ECE Secretariat with a vision on an integrated approach on air pollution, i.e. integrating environmental and economic aspects of mitigation of air pollution, liked the model and recommended to expand RAINS with a sub-model that could address economic issues. In 1985 representatives of UN-ECE attended a workshop at IIASA which simultaneously served as a scientific review of the RAINS model and a workshop for policy makers. It gave the attending policy makers an impression of the scientific credibility of RAINS. It also gave them the possibility to do suggestions to improve the possibilities of the model to use it as a tool in policy making. See also Farrell and Keating (2005) and VanDeveer (2005). Several suggestions were made by policy makers at that workshop which were later implemented in RAINS. This proved to be crucial for the relevance of RAINS later in the CLRTAP process. One of the recommendations was to improve flexibility of the model: negotiators of countries present at the workshop pointed out that “politicians need a menu to choose from: not a recipe”. In other words: they made clear that policy makers need to be able to make the decisions themselves. The possibility of choice and sufficient alternatives for negotiations were important for the negotiators. Thus at the workshop policy makers specified what they saw as their own role and what the role of scientists should be. This paved the way for an enhanced legitimacy of use of the models in the process.

4.2.2.2. Objects. In the same year UN-ECE arranged a channel in which RAINS could be used. The Group of Experts on Costs and Benefit Analysis, at its second meeting in 1985 considered that “the IIASA model (i.e. RAINS) could provide a useful tool for cost-effectiveness analysis” (UN-ECE, 1985). In 1986 the Task Force on Integrated Assessment Modelling (TFIAM) was established in which RAINS and other models would be used more and more intensely in the years that followed. In this process the models were tuned towards the needs of the negotiators. The further development of the models became policy-driven, but at the same time stayed in touch with latest scientific developments, thus carefully balancing between policy relevance and scientific credibility of the model. See for more discussions of the RAINS model as a boundary object Sundqvist et al. (2002) and Cash et al. (2003).

4.2.2.3. People. In this early phase halfway the 80s, certain individual people, through their specific qualities, interventions and visions, have been playing an important boundary role. The representatives of for example the RAINS-team, but also the EMEP-modelling team and the developers of the critical load concept, enhanced simultaneously credibility, legitimacy and relevance through the way they could present the possibilities of the model, and the way they could take policy considerations into account in new model formulations and concepts. See also VanDeveer (2005). At the same time, by doing so they played an important role to provide new framings for the policy problem and enhance discussions and negotiations. In other words, they did not merely communicate “science results to policy makers” or played the role of scientists offering a listening ear to policy makers, but played an important role in the policy problem framing itself. As important was the ability of those individuals to maintain credibility in the science domain.

Also individual policy makers, negotiators in country delegations like the Dutch delegation mentioned above, and civil servants like the people in the UN-ECE secretariat mentioned above, played an important role. Not only did they identify specific tasks for the modelling groups, like asking for menus instead of recipes, but also they enhanced model development by being visionary about possible applications like proposing to add economic modelling to make the modelling work more integrated. Thus, they played an important role in enhancing legitimacy and relevance of the models.

4.3. Participation

4.3.1. Participation in an international data collection effort

The number of science and policy actors participating in the CLRTAP process is large. One of the reasons for this
large network of actors is the need for data collection. The signatories of CLRTAP are responsible for the provision of the data from their own country. Data are needed for e.g. the calculation of atmospheric transport of pollutants in the EMEP models, for emission inventories and for the compilation of maps with critical loads, all of which are integrated in the modelling work performed by TFIAM. Furthermore, emission inventories are needed for the monitoring of compliance. The fact that countries are responsible for their own data contributes to the credibility and legitimacy of the modelling work of EMEP and TFIAM.

4.3.1.1. Objects. One element of the data collection within CLRTAP is the standardisation of the methods to collect data. Different methods for data collection exist, within CLRTAP certain choices have been made and agreement has been reached on the appropriate way of collecting data in this setting. See for example the EMEP/CORINAIR Guidebook (EEA, 2001) and the UN-ECE Mapping Manual (UN-ECE, 2004). This included several compromises and simplifications. Gough et al. (1998) describe how data used in CLRTAP are in fact “negotiated” data, (see also Tuinstra et al., 1999). Whether those data belong to the science domain or the policy domain depends at what stage of the process of the data collection they are considered. The development of the EMEP/CORINAIR Guidebook is seen as a scientific activity. The collection of the data is in most countries done by national scientific institutions. However, at the moment that the data are used as an input for model calculation within TFIAM, the modellers label the datasets as being the input from the countries, and thus being political. In the modellers view the international data collection effort, which facilitates participation of all the countries, enhances rather the legitimacy than the credibility of their work.

4.3.2. Participation in TFIAM

Participation in the TFIAM gradually changed during the years of its existence. First it was a forum for different modelling groups for comparing and discussing different modelling approaches. As such, it functioned as a kind of peer review group for experts mainly meant to build credibility for the modelling work. Important for this credibility was, that in the beginning not one but various models were used to perform similar calculations. Currently, RAINS is the main model used. The main goal of TFIAM is now to advice on the development of RAINS.

4.3.2.1. People. Over the years, TFIAM has grown and the current participants do not necessarily represent groups that contribute to the modelling work in TFIAM itself. Rather they represent countries. Participants have the possibility to give presentations of national modelling efforts and in this way TFIAM remains being the forum to be kept up to date with the latest modelling developments. However the active participation of the different national experts also serves another goal. It can avoid a feeling among the parties that the RAINS model and what happens in TFIAM is a “black box” on which the countries can have no control. Credibility and legitimacy of RAINS increases when representatives of countries have the capacity to know what is going on and can explain the “black box” to their governments. In doing so, the participants in TFIAM play a boundary role. At the same time it should be noted that some countries within UN-ECE are more active than others and some countries do feel more involved than others. Castells and Nijkamp (1998) elaborated on this point for southern European countries. See for more details about the participation of eastern European countries Botcheva (1998) and VanDeveer (2005).
legitimacy which however was seen as very important for consensus and continuation.

As mentioned above it is important to note, that initially more integrated assessment models were considered in TFIAM than RAINS only. Current literature frequently overlooks this fact. The availability of three integrated assessment models in CLRTAP played an important role in building credibility and legitimacy (Gough et al., 1998; Tuinstra et al., 1999). Jäger (1998) also notes that using multiple models with different approaches to the problem in integrated assessment processes can enhance the credibility and legitimacy of the processes. Currently there are two tendencies within TFIAM with regard to the use of multiple models. On the one hand there is the tendency to avoid “duplication of work” and to stress the consensus already reached on the use of certain models, like the atmospheric model EMEP, and the integrated assessment model RAINS. But there is another tendency as well: the development of ensemble modelling, meaning comparing and analysing the outcomes of different models in calculating the same phenomena. This is being done for atmospheric modules in the context of the work for the European Commission. For economic modules this is not being done, apparently because approaches in various models available are comparable. Apparently in different scientific domains there are different views of the amount of possible approaches to a problem.

4.5. Dissent

4.5.1. Dealing with dissent through consensus documents and common concepts

4.5.1.1. Texts. Above we already discussed the importance of boundary texts to establish credibility for the use of the RAINS model. In general in CLRTAP a very important role can be attributed to the reports produced of the various meetings of working groups and task forces. Those are consensus documents which form together the collective memory of CLRTAP and ensure credibility and, even more, legitimacy of the work. That the content of those reports and the procedure leading to the report have not been going undisputed can be illustrated by the correspondence between the TFIAM chairman and one of the UK-delegates in TFIAM about a crucial meeting in 1993 in the last phase of the preparations for the Oslo Protocol. The UK-delegate challenged the wording in the document with regard to the use of one model and also expressed his concern about the way the document was produced, not reflecting the discussions, but only final conclusions. This led to an in-depth reaction of the chairman in which he not only discussed the procedure leading to the report and the actual wording but in which he also sketched what he considered the appropriate way of working of the TFIAM and the supposed behaviour of delegates. He argued that the TFIAM was not a political body but a scientific one and that the way of operating of the UK-delegation was not according to that. He also argued that participants in the TFIAM, different than in WGS, are in fact not considered “delegates”, but experts not representing any country. This example illustrates the apparent importance of the TFIAM reports as boundary texts and the reactions which are provoked when challenging those texts.

4.5.1.2. Objects. At the 7th meeting of WGS in 1992 the gap-closure concept was proposed by the Norwegian delegation. This concept implies that “the gap” between current levels of atmospheric deposition and the critical loads is “closed”. The introduction of the concept happened when it was difficult to reach an agreement on the kind of targets for ecological protection to base further negotiations on. It was accepted by all parties and was a breakthrough in both the science and policy processes in the preparatory phase of the Oslo Protocol. It was appealing because it formed a direct link to critical loads in each grid-cell, but also implied a kind of equity, because the percentage for closing the gap is the same everywhere. The concept offered a way to deal with dissent and formed an important point of departure to reach consensus. Finally, the negotiators chose a “60% gap-closure” scenario as a starting point for the negotiations, in the 10th WGS meeting 1993. Interestingly, the concept was proposed in the WGS, which is a “political” body under the Convention, and not, which one could have expected, in TFIAM, which is a “scientific” body. However, it was a scientific expert in the Norwegian delegation who made the proposal, a meteorologist. This is also interesting because this kind of concept one would expect from an economist rather than form a meteorologist. Looking more closely, we see that though in this meeting the Norwegian expert acted being part of the Norwegian delegation, in the Convention he also played an important role in the scientific work of the Convention, because he was the director of the institute that developed the EMEP model.

5. Discussion and conclusions

The main question of this article was what role science–policy interaction or boundary work plays in shaping assessment frameworks and assessment processes within CLRTAP. Also we asked how participants in the CLRTAP assessment process divided and co-ordinated work between science and policy and how this enhanced credibility, legitimacy and relevance with multiple audiences. The application of a combined framework for analysing credibility, legitimacy and relevance and for analysing boundary work turned out to be helpful in describing in detail what happens in practice at the science–policy interface. Overall, we can conclude from the analysis of the CLRTAP assessment process that credibility, legitimacy and relevance to many audiences were enhanced by boundary work in an early stage of
the process. Furthermore we conclude that boundary work in the assessment processes in CLRTAP has been enhanced because a forum was provided, where boundaries between science and policy could be discussed. This enabled a successful division and co-ordination of work between science and policy which made the boundaries in the CLRTAP process to remain quite stable through the years. Finally we conclude that this has as implication that new boundary work is needed when knowledge produced in the context of CLRTAP is used in other policy contexts or when new participants enter the arena.

5.1. Boundary work and design elements in the CLRTAP assessment process

Our analysis of initiation and goal of the CLRTAP and the participation in CLRTAP shows that what is considered to be credible, legitimate and relevant is established already in an early stage of the development of the assessment framework. It is therefore important for actors to be involved in boundary work in an early stage of the communication process. As we have seen, for example members of the RAINS team were very early participants in boundary work. The features of the RAINS model currently match quite well with what is considered relevant in the CLRTAP community. RAINS, the policy development within CLRTAP and the set up of the data collection structure developed in parallel and influenced the course of each others developments. RAINS clearly participated in setting the scene and could therefore enhance its own relevance. An example of this is the development of the use of the concept of Critical Loads, which could not have been operationalised without integrated assessment models such as RAINS and the other models used.

With regard to dealing with uncertainties we also see that this requires careful boundary work balancing between credibility and legitimacy eventually leading to operational structures and effective assessment procedures. The example of the use of models and monitoring equipment which could be managed and applied in all countries (“lowest common denominator standard”) shows that this balance is of crucial importance for continuation of e.g. the monitoring programme without which this programme never could have been effective for policy making. The context determines what kind and degree of uncertainty is being accepted. With regard to dealing with dissent we see that the consensus structure of CLRTAP, which is inherent to the way of working of the UN-CEC, offers ample room for boundary work. We showed this in the example of the production of consensus reports but also in the creation of the “gap-closure” concept which was inspired by the necessity to come to consensus in the policy debate. The concept served a boundary role by being relevant through its ample timing and in fitting both to the policy concerns and the framing of the scientific debate at that stage. In its context it was both to multiple audiences politically legitimate because of its equity dimension, and scientifically credible because of its connection to environmental effects.

5.2. Co-production of science and policy in CLRTAP

The assessment process within CLRTAP is an interesting illustration of co-production of air quality science and air quality policy in Europe. The development of e.g. the concept of critical loads as an important example has been described elsewhere (Bäckstrand, 2001; Sundqvist et al., 2002; Grennfelt and Hov, 2005). The example we highlight here is the establishment of the TFIAM and the iterations between the RAINS team and national negotiators and UN-ECE staff. They played an important role in the further development of the model. Simultaneously the possibilities of the model itself inspired the extension of the then existing CLRTAP structure with another expert body (TFIAM) and the involvement of other integrated assessment models. In the course of the years the importance of TFIAM increased in CLRTAP because of its integrative function in the science–policy network. The air pollution problem was further framed through the tools and data available. The fact that TFIAM with the help of integrated assessment models was able to produce quantitative scenarios for alternative policy decisions helped the CLRTAP process to move forward.

In this light it is not surprising that the outcome of the RAINS-review in 2004 has been positive in the sense that it judged the model “sufficiently credible to be used as a tool in policy making” (Grennfelt et al., 2004). The terms of reference for the review have been developed in the same social structure and knowledge frame as the structure and knowledge frame that developed RAINS. The evaluation was positive because it focused on exactly those issues that were found important by the community which themselves had produced the Integrated Assessment framework of which RAINS is a part. It is important to note that this does not mean that the review has not been scientifically correct or credible itself. Rather it shows that apparently certain boundaries are quite stable. What is expected from science within policy, and the conviction on what in this setting is credible, legitimate and relevant, has remained quite stable within the issue domain of air pollution in Europe.

5.3. Moving boundaries

An issue strongly related to co-production and the importance of being involved in an early stage in boundary work is transparency of the process. For “newcomers” transparency may be lower than for those involved from the beginning. The language used and rules of the game might have developed and established slowly to a kind of jargon within the existing group. What will be discussed, what is relevant, what is the “language” spoken and what
scientific disciplines are supposed to be relevant has already been established. Together with the knowledge and the framing of the problem also the social structure for providing that knowledge has been established. For those who were not participants in this co-production process this might be difficult to change. Ideas about what should be the tasks of science and policy might differ for newcomers, because they were not involved in the early boundary work. Thus they will have a different view on credibility and legitimacy because they will have different expectations of scientific and policy actors, and on relevance, because they have a different problem frame of the assessment.

This observation regarding different views on credibility, legitimacy and relevance is of importance because (1) it has consequences for the transferability of the knowledge produced within CLRTAP into other arenas and (2) when more newcomers enter the arena, credibility, legitimacy and relevance will be contested and new boundary work and adaptation of knowledge frames is needed. Currently, knowledge produced within the CLRTAP arena, indeed is being transposed into another arena, namely the work of the European Commission. It will be interesting to follow this development further. See also for an analysis of the linkages between CLRTAP and European Commission work Wettstead (2002) and Selin and VanDeveer (2003).

The actors within CLRTAP are currently confronted with the question of the future and the relevance of the Convention itself. What role can it play next to EU policies? What is the relevance of the air pollution issue in broader contexts? Is it possible to extend to other issue areas like climate change and hemispheric (global) transport of air pollutants? What is the effect of increased stakeholder involvement? All those issues ask for new stages of boundary work. Early boundary work is important but boundaries also move when new participants enter and when contexts change. It shows that views on credibility, legitimacy and relevance of assessments change with moving boundaries and that credibility, legitimacy and relevance have to be worked upon continuously.

These are also an important lesson for other policy areas. In environmental policy, be it on a local, European, or global scale, problem framing and knowledge-needs are not a given. We see this e.g. in the field of climate change. In time, problem framings of the climate issue have been ranging from a carbon dioxide problem and an energy problem to an issue of global environmental change in general and sustainable development to an issue of human security and adaptation. These differing problem frames imply differing knowledge-needs. Other scales apply and other data and computer models are relevant then previously. In addition, not only knowledge-needs change, also the relevant policy settings, timing of decisions and relevant stakeholders to deal with the issue change. On a regional scale, the climate issue in e.g. the Netherlands is now being connected to spatial planning and water policies (Kabat et al., 2005). This involves other policy makers and stakeholders then those previously involved, like e.g. water boards and regional planning authorities. It also asks for a different kind of scientific expertise and application of scientific knowledge. Adapting assessment designs, careful communication between the various relevant policy and science actors and awareness of the context will enhance credibility, legitimacy and relevance of assessments in such changing settings.

Acknowledgements

This research has been made possible by funding of the Environmental Systems Analysis Group, Wageningen UR and the Netherlands Environmental Assessment Agency, Bilthoven.

Furthermore the authors wish to express thanks to the scholars and fellows from Harvard University’s Global Environmental Assessment Project and the fellows of the “Rethinking Project” led by Twente University (the Netherlands), for sharing valuable thoughts. Peter Buïjtjes, Willem Kakebeeke, Rob Maas and Arthur Petersen and three anonymous reviewers provided valuable comments on earlier versions of this paper. Any remaining errors, however, remain the responsibility of the authors.

References


Kennedy School of Government, Harvard University, Cambridge, MA.


Halffman, W., 2003. Boundaries of Regulatory Science: Eco/toxicology
Guston, D.H., 2001. Boundary organizations in environmental policy and
Grennfelt, P., Erisman, J.W., Forsberg, B., Fowler, D., Fudala, J., Hov,
Jasanoff, S. (Ed.), 2004. States of Knowledge: The Co-Production of
CSCE, 1975. Final Act, Chapter II, Section 5. Conference on Security and
Co-operation in Europe. Helsinki, Finland.
Eckley, N., 2001. Designing effective assessments: the role of participa-
tion, science and governance, and focus. Environmental Issue Report
26, EEA Copenhagen.
assessment processes in consensus negotiations. Global Environmental
Change 12, 15–23.
sulfur to POPs. In: Farrell, A.E., Jäger, J. (Eds.), Assessments of regional
and global environmental risks: designing processes for the effective use
of science in decision making. RFF Press, Washington, DC.
EEA, Copenhagen. Updated. reports.eea.eu.int/EMEPCORINAIR4.
Farrell, A.E., Jäger, J. (Eds.), 2005. Assessments of Regional and Global
Environmental Risks: Designing Processes for the Effective Use of
Science in Decision making. RFF Press, Washington, DC.
Farrell, A.E., Keating, T.J., 2005. Dissent and trust in multilateral
assessments’. In: Farrell, A.E., Jäger, J. (Eds.), Assessments of Regional
and Global Environmental Risks: Designing Processes for the Effective
Use of Science in Decision Making. RFF Press, Washington, DC.
ments: four underappreciated elements of design. Global Environ-
mental Change 11, 311–333.
Pinch, T., Markle, G.E. (Eds.), Handbook of Science and Technology
emerging methodology for complex issues. Environmental Modelling
and Assessment 3, 19-29.
Grennfelt, P., Erisman, J.W., Forsberg, B., Fowler, D., Fudala, J., Hov,
Report, Service Contract for Review of the RAINS Integrated
Assessment Model—Reference ENV.C1/SER/2003/0079; http://
europa.eu.int/environment/air/cafe/activities/rains_model.pdf.
Grennfelt, P., Hov, Ö., 2005. Regional air pollution at a turning point.
Ambio 34, 2-10.
Environmental Change Through Knowledge and Negotiation. Lessons
from the Rhine Action Program and the Second Sulphur Protocol,
Guston, D.H., 2001. Boundary organizations in environmental policy and
Halfman, W., 2003. Boundaries of Regulatory Science: Eco/toxicology
and Aquatic Hazards of Chemicals in the US, England and the
Netherlands. Dissertation, Science Dynamics, University of Amster-
dam, pp. 35–72.
use of critical loads in emissions reduction agreements in Europe.
Water Air Soil Pollution 85, 2381–2388.
Hordijk, L., 1991. Use of the RAINS model in acid rain negotiations in
Jäger, J., 1998. Current thinking on using scientific findings in environ-
mental policy making. Environmental Modeling and Assessment 3,
143–153.
Harvard University Press, Cambridge, MA.
Jasanoff, S., 1996. Beyond epistemology: relativism and engagement in
the politics of science. Social Studies of Science 26, 393–418.
Jasanoff, S. (Ed.), 2004. States of Knowledge: The Co-Production of
Science and Social Order. Routledge, London.
Jasanoff, S., Wynne, B., 1998. Science and decision making. In: Rayner,
S., Malone, E.L. (Eds.), Human Choice and Climate Change. Battelle
Press, Columbus, OH.
Fruits of the cold war: the convention and the first sulphur protocol.
In: Sliggers, J., Kakebeeke, W. (Eds.), Clearing the Air: 25 years of
Lidskog, R., Sundqvist, G., 2004. From consensus to credibility: policy-
relevant science in late modernity, innovation: the European. Journal
of Social Science Research 17, 205-226.
Lövblad, G., Tarrasón, L., Tørseth, K., Dutchak, S., 2004. EMEP
Assessment. Part I. European Perspective. EMEP, Oslo.
Integrated assessment modelling—the tool. In: Sliggers, J., Kakebeeke,
W. (Eds.), Clearing the Air: 25 years of the Convention on Long-
Range Transboundary Air Pollution. United Nations Economic
Miller, C., 2001. Hybrid management: boundary organizations, science
policy, and environmental governance in the climate regime. Science,
Technology and Human Values 26 (4), 478–500.
www.ksg.harvard.edu/geo/pubs98sw_boundaries.pdf.
Mitchell, R., Clark, W., Cash, D., Dickson, N. (Eds.), in press. Global
Environmental Assessments: Information and Influence. MIT Press,
Cambridge, MA.
Nilsson, J., Grennfelt, P. (Eds.), 1988. Critical Loads for Sulphur and
Nitrogen. Report 1988:15. Nordic Council of Ministers, Copenhagen,
Denmark.
Nordberg, L., Bull, K., Graham, O., Jagusiewicz, A., Sand, P.H., Tollan,
A., Wuester, H., 2004. The role of the secretariat: building the protocol
tree. In: Sliggers, J., Kakebeeke, W. (Eds.), Clearing the Air. 25 Years
of the Convention on Long-Range Transboundary Air Pollution,
Policy Studies Review 16, 104–137.
Rayner, S. (Ed.), 2006. What drives environmental policy? Editorial,
Global Environmental Change 16, 4–6.
Schneider, T., Schneider, J., 2004. EMEP — backbone of the convention.
In: Sliggers, J., Kakebeeke, W. (Eds.), Clearing the Air. 25 Years
of the Convention on Long-Range Transboundary Air Pollution.
European air pollution politics. Global Environmental Politics 3,
construction of an emerging policy field. Science and Public Policy
Sliggers, J., Kakebeeke, W. (Eds.), 2004. Clearing the Air. 25 Years
of the Convention on Long-Range Transboundary Air Pollution.
pollution abatement strategies. Environmental Science & Policy 5,
147–156.
Tuinstra, W., Hordijk, L., Amann, M., 1999. Using computer models in
international negotiations/the case of acidification in Europe. Environ-
ment 41, 33–42.
Human Environment, Stockholm.
UN-ECE, 2004. Manual on methodologies and criteria for Modelling and
Mapping Critical Loads & Levels and Air Pollution Effects, Risks and
Trends; www.icpmapping.org.

VanDeveer, S., 2005. European politics with a scientific face. Framing, asymmetrical participation, and capacity in LRTAP. In: Farrell, A.E., Jäger, J. (Eds.), Assessments of Regional and Global Environmental Risks: Designing Processes for the Effective Use of Science in Decision making. RFF Press, Washington, DC.