EVALUATING THE EFFECTIVENESS OF ENVIRONMENTAL POLICY:
AN ANALYSIS OF CONCEPTUAL AND METHODOLOGICAL ISSUES

Joos Gysen
Catholic University of Leuven
Belgium
email: Joos.Gysen@hiva.kuleuven.ac.be

Kris Bachus
Catholic University Leuven
Belgium
email: Kris.Bachus@hiva.kuleuven.ac.be

dr. Hans Bruyninckx
Wageningen University
The Netherlands
email: Hans.Bruyninckx@alg.swg.wau.nl

Paper presented at the European Evaluation Society Seville Conference
October 10-12, 2002, Seville, Spain

Please do not quote without permission from the authors
Abstract

Problems of implementation and effectiveness have recently been subject to growing interest in the field of policy evaluation, in particular at the EU level. The literature as well as the praxis in this field have also highlighted a number of problems, which can be divided into three categories.

The first type are conceptual problems; one example can be found in the definition of the concept of effectiveness. The second type are the methodology-based problems, such as the causality question. Finally, important empirical obstacles exist, mostly related to the (non-) existence of input and output indicators and the quality of data.

In our paper we will analyse these problems and their consequences for the field of environmental policy evaluation.

The findings of the paper are based on the first results of the work of the Centre of Expertise for Environmental Policy Sciences (CEEPS), funded by the Flemish government for a period of five years (2002-2007).
1. Introduction

This paper aims to address some central issues relating to effectiveness evaluation in the field of environmental policy.

Recently, implementation and effectiveness of policy has been subject of inquiry at different policy levels and in different policy fields. At the European level the European Environmental Agency (EEA) has been very active trying to answer the question: “are we being effective” (EEA: 2001). In order to actually provide an answer to this question, the EEA is aiming to develop a methodology, which counters the difficulties that come along with this question. Besides the EU, other institutions are active on this issue: scholars and practitioners (e.g. administrations) alike have been engaged in answering the effectiveness question. Most initiatives have been designed around specific environmental subfields, e.g. the evaluation of specific water treatment policies, air pollution abatement policies, etc. In most instances, a case study forms the fundament of the research and evaluation methodology. With this paper we wish to contribute to a number of efforts that provide elements for more universally applicable methodologies for measuring the effectiveness of environmental policy (Mickwitz, 2000 Bressers, 1991).

We emphasise some conceptual and methodological issues within the field of environmental policy evaluation, not only because of our personal interest in specific methodological and conceptual challenges in this subfield, but also because environmental policy evaluation is of increasing importance. Environmental issues have achieved a ‘sense of urgency and seriousness’. They have been accompanied by an increase in both public and political interest. In addition, the amount of financial and other resources attributed to both environmental policy making and environmental policy evaluation has increased dramatically. Yet, environmental policy is a relatively ‘young’ policy field. The first serious policy initiatives date from only a few decades ago. In fact, in Flanders, environmental policy has been a substantive and meaningful competence of the region, only since the constitutional reforms of 19801. This makes environmental policy in Flanders, indeed, a ‘very young’ policy domain. The evaluation of both new and older initiatives has only just started in the last several years. The establishment of the Centre of Expertise for Environmental Policy Sciences (CEEPS), and its task to develop policy evaluation instruments, has to been placed in that framework.

In this paper we will address the central concepts of effectiveness and causality and their role in environmental policy evaluation. We will do this in light of some of the key characteristics of environmental policy. We will end with a methodology for evaluating environmental policy, which will focus on measuring effectiveness of environmental policy.

2. The concept of effectiveness

Effectiveness takes a central position in policy evaluation. The dominant question in effectiveness evaluation is ‘does this policy work?’ This is an inherently normative question because it involves an appreciation of what is meant by ‘work’. However, this is not the only evaluation question to be posed. The EEA2 identifies three basic categories of evaluation questions:

- Firstly, questions can be descriptive. These questions try to address the ‘what has happened?’ question.

---

1 Environmental issues in Belgium are largely the responsibility of the three regions, namely Flanders, Wallonia and Brussels. The Federal government has a number of limited specific competences. However, the international dimension of Belgium’s environmental policy is still the responsibility of the Belgian government, this means that a common point of view has to be reached and defended. The defederalisation of environmental policy making also means that large differences exist between the three regions in terms of legal framework, instruments, and results.

- Secondly, questions can be of causal origin. Causal questions do not simply ask ‘what has happened?’; they also try to attribute the happening, the change, to the policy.

- Thirdly, questions can be normative. These questions evolve around the central question of satisfaction, ‘Are the results satisfactory?’ According to the EEA, this question can be answered within the perspectives of effectiveness, relevance, efficiency or utility. We are especially interested in the effectiveness perspective. We do realise the importance of relevance, efficiency and utility but the scope of this paper is clearly ‘measuring the effectiveness of environmental policy’.

As will become apparent in the remainder of this paper, we believe that the effectiveness question cannot be separated from the two first types of evaluation questions (what has happened? and the causality question), if it is to be answered in a realistic and satisfying way. The paper strongly emphasises the fact that effectiveness evaluation in the case of environmental issues builds on well-documented and developed answers to the descriptive and causal questions.

To further refine the conceptual framework, we will make a distinction between four kinds of effectiveness: institutional effectiveness, target group effectiveness, environmental effectiveness and societal effectiveness (see further). This distinction will prove to be important in the light of the evaluation methodology in general and more specifically it will indicate differences in the conceptualisation of causality.

Closely related to effectiveness is the issue of effects. Which types of effects do we take into account to measure effectiveness in the light of a stated set of policy objectives? This question seems rather straightforward or even trivial. However, for many policy evaluations it is not. The evaluation of the environmental voluntary agreement (EVA) between the Flemish government and the local communities and provinces in Flanders (Bachus et al., 2001) serves as an example for the lack of attention to policy effects. Information on effects of the EVA were hardly monitored or identified. As a result, effects were very hard to measure in function of a policy evaluation, since no precaution was taken to monitor them. An effectiveness appreciation of the instrument was therefore very difficult.

Policy effects can be divided in three major categories: output effect, outcome effect and impact effects. Further in the paper, we will illustrate these different effects by an example in the field of water policy. The different effects also adhere primarily to the different types of effectiveness. In the next paragraphs we will link the different types of effectiveness to different effects. To illustrate the place of the various effectiveness types and effects we will use figure 1, which is an adaptation of the EEA (2001) policy evaluation frame.

The basis of the model are two interconnected ‘environments’. On one side we have the policy process. On the other, what the EEA calls the ‘outside world’. Clearly, the outside world provides for input and interaction on every stage of the policy process. However, we will limit ourselves to the most important inputs. The policy process comprises of three parts, the formulation of objectives, the choice and use of instruments and the output. The policy process gets input from, and it produces effects for, the outside world. The outside world, through the expression of societal needs gives the input for the policy objectives. Also, societal means are input factors in every step of the policy process. The policy process produces output. This output translates into outcome and impact, both effects in the ‘outside world’.

---

3 This evaluation was the first full scale evaluation of an environmental policy instrument in Flandres.

4 We regret the use of this term. By using ‘outside world’ the EEA draws a too thick a line between the inner and outer dimensions of policy. Because we make use of the EEA framework as a starting point we will, however, maintain the term.
In order to fully grasp the model we need to further elaborate on different types of effects and effectiveness.

### 2.1 Institutional effectiveness

Bruyninckx and Cioppa (2000) refer to institutional effectiveness as ‘suggesting that a regime [or policy] is operating in some sort of agreed-upon fashion [...]’. In other words and translated to policy: ‘institutional effectiveness’ is the extent to which the output of the policy matches the objectives of the policy. The output is defined as the tangible results of a measure (EEA, 2001). Output has a rather short-term dimension. Outputs can be noticed shortly after the implementation process of the policy instrument or even during the process, e.g. when a new type of permits becomes compulsory, the amount of permits granted from the first day will be an output effect. These output effects are not necessarily part of the actual policy objectives. They do not necessarily have an automatic or direct relationship with the policy performance. For example, in the field of environmental policy, the number of environmental permits granted, or the number of hired field inspectors, can be policy objectives, but they do not give a direct indication that the environment is improving. They can only give an indirect indication for the intended policy results. However, in practice, these output effects are often the only effects that are monitored.
2.2 Target group effectiveness

Secondly we distinguish what we call ‘target group effectiveness’. By this we mean the degree to which the outcome, defined as the response of the target groups to the output (EEA, 2001) of the policy corresponds with the policy objectives. Where the output effects can take place in the short term, outcome effects are most likely to occur in the middle long term. We use a broad conceptualisation of target group response or behaviour. Behaviour can be anything going from individual behaviour, over group behaviour, to societal activities. For example in the field of environmental policy, the change in environmental infringements, is a possible outcome of the number of hired field inspectors. Other examples of environmental policy outcome effects are reductions in emissions, increased recycling rates, shifts in the use of transport modes.

2.3 Impact effectiveness

A third type of effectiveness is what we refer to as ‘impact effectiveness’. We refer to impact effectiveness to emphasise the policy issue as such. We define any effect that influences the state of the policy issue as an impact effect. Mickwitz (2000) would call this effect the ‘ultimate outcome’. We choose not to use this term because of the significant distinction we make between outcome and impact. In the field of environmental policy we would use the term ‘ecological or environmental effectiveness’. Bruyninckx and Cioppa (2000) describe this effectiveness by using the case of environmentally effective institutions:

‘An environmentally effective institution eradicates or alleviates anthropogenic deductions from and/or deposits to an ecological system or systems in balance with the system's natural regenerative processes.’

In this definition the antropogenic input (deposits) and outputs (deductions) are framed in a natural equilibrium. This definition provides for a basis for evaluation, independent from the immanent objectives of the institution or policy. Defining environmentally effective policy in this way provides a basis for policy evaluation, stretched to its most comprehensive form. Evaluating the policy’s contribution to the natural regenerative balance of natural is using the fundamental bottom-line. In other words, this bottom-line goes further then the environmental objectives which might be stated in the policy.

The impact effects or impact (EEA, 2001) of a policy is often only visible in the long term. In the example of environmental policy, impact effects can be categorised as part of the state of the environment. Often these effects are expressed in terms of quality. A better air quality could be a possible impact effect if the policy was aimed at reducing the unauthorised emission of air pollutants.

2.4 Societal effectiveness

Lastly we identify ‘societal effectiveness’. Within the EEA framework, societal effectiveness corresponds with both the relevance and utility question. Societal effectiveness addresses the question whether or not the impact (or impact effects) satisfies the societal needs. In other words: ‘is the effect a contribution to broader societal objectives?’. In the case of environmental policy, the dominant policy discourse or framework is sustainable development for the moment. This means that the effects can be tested on their sustainability calibre.

---

5 European Commission, Towards a methodology for evaluating the effects of measures taken to implement EU environmental legislation, first draft.
2.5 Illustrative example

The policy objective of a hypothetical water policy is clean(er) water. The appointed instrument to attain this objective is building water treatment plants. One of the outputs of this policy will be the number or increased capacity of operational water treatment plants. An aspired and possible outcome will be the amount of treated, contaminated water in these plants. The behavioural change is the going from (still hypothetical) ‘not treating water’ to ‘treating a certain amount’ of water, which as we observe in the Flemish case has fundamental implication for policy makers and actors at different levels in many ways. The impact of this policy is the improved water quality. Figure 2 illustrates how objective, instrument and the different effects are interconnected.

Figure 2 Hypothetical water policy: objective, instruments and effects

In order to measure the impact effectiveness we compare the quality of the water (impact) with the stated policy objective as far as they set any water quality standards. If we were to measure the target group effectiveness we would have to look for objectives within the policy with regards to, for this example, the ‘treatment of water’ behaviour or in other words for some measure in the objectives which states the amount of treated water that is aimed for. To measure the output effectiveness, we need to compare the output with the output objectives. In this case we would have to compare the number or capacity of operational water treatment plants with the number of operational plants, which the policy aims at. Lastly, we can evaluate the societal effectiveness by checking the coherence of the policy (in all its elements) with sustainable water management objectives formulated by for example the EU in its water directives.

Above we have defined the different types of effects and effectiveness. In figure 1 we indicate which place they take in policy evaluation. In the next paragraphs we will further elaborate on two important concepts in environmental effectiveness evaluation which appear in the figure: objectives, instruments.

2.6 Objectives

We have used the term policy objectives in a sort of evident way. Experience with environmental policy evaluation demonstrates, however, that for evaluation purposes the term is often rather problematic, if not in an essential way, than certainly from an evaluator’s more pragmatic point of view. The use of the term “policy objectives” refers to the official policy objectives. These objectives are made explicit through documents directly related to, in this case, environmental policy. However, objectives can also be found in documents, which are not directly related to the environmental policy under review. It is possible that other policy fields, such as public health or town and spatial planning, provide some objectives that are strongly related to the policy that is being evaluated. These objectives can be environmentally oriented. They can
also be policy objectives with relevance to environmental policy without being environmental objectives. An example of this group of objectives is the number of illnesses, which are directly related to some sort of pollution, for example in the case of carcinogenic substances. In order to attain this non-environmental objective, some environmental measures need to be taken with regard to the source of the disease, in casu the pollution.6

Because of the relevance of indirect policy objectives, we define the policy objectives for the purposes of evaluating a policy as ‘the objectives which are made explicit through official policy texts. These policy texts can be directly or indirectly related to the policy’. It is obvious that careful consideration has to be given to the inclusion of indirect or secondary policy objectives. Depending on the parameters of the evaluation exercise they will play a prominent role or will be considered as trivial. In any case it is useful, if not necessary, as an evaluator to be aware of this.

2.7 Instruments

Environmental policy evaluations are often linked to the use of a specific instrument or set of instruments. An effectiveness evaluation is in that sense also an evaluation of those instruments in the policy context in which they are used. It is therefore important to have a good understanding of the concept. In the EEA frame, what we call instruments, is indicated as input. We have chosen to use the term ‘instrument’ in stead of input, to stress the importance of the instrument in the policy process. The instrument sensu strictu is accompanied by other components (inputfactors according to the EEA concepts, EEA 2001): staff, administrative structures, financial means, training, awareness raising. In other words the resources dedicated to the design and implementation of a measure. Hoogerwerf (1989) tends to follow this broad conceptualisation by defining instruments as everything that is used by or through a government to attain the policy objectives. Bressers (in Glasbergen 1994) on the other hand, limits this definition by excluding elements such as planning procedures, auxiliary resources, money and manpower.

Other authors add a more explicit reference to the objectives of using certain instruments in general. Vedung (1998) stresses the finality of policy instruments by defining them as “the set of techniques by which governmental authorities wield their power in attempting to ensure support or prevent social change”. Mickwitz (2000) combines Vedung’s definition of policy instruments with Lundqvist’s (1996) ‘purpose definition’ of environmental policy – “courses of action which are intended to affect society – in terms of values and beliefs, actions and organisation – in such a way as to improve, or prevent the deterioration of, the quality of the natural environment”. This results into a broad definition of environmental policy instruments: “the set of techniques by which governmental authorities wield their power in attempting to affect society – in terms of values and beliefs, actions and organisation – in such a way as to improve, or prevent the deterioration of, the quality of the natural environment”.

Environmental policy instruments – as well as policy instruments in general – can be categorised in different ways. We follow Van der Doelen’s (1993) division of policy instruments in three major groups. The first is the group of economic measures such as subsidies or taxes. The second is the group of legal instruments like laws and permits. Finally Van der Doelen works with a group called communicative instruments. Examples of this group are sensitising and education. The OECD (1994) and Vedung (1998) applied this categorisation to environmental policy instruments by translating Van der Doelen’s legal instruments to ‘regulation’ and by translating communicative instruments to ‘[per]suasive instruments’.

The importance of the choice of instrument is emphasised by Bressers (1991) when he states that the choice of the policy instrument influences the course of many processes in the policy field. Bressers (1994) continues by stating that the effectiveness of instruments is partly determined by the instrument mix in which the instruments evolve. He lists four issues with importance for effectiveness evaluation.

6 We acknowledge that other sources, beside pollution, can exist.
Dynamic of time. Not all actors react in the same way. This will influence the behaviour of the target group in time.

Complex organisations. He defines target groups as not merely individuals. Often they are organisations. These organisations are not homogeneous. The point of entry can therefore influence the effectiveness.

Environmental behaviour of the target group and environmental effects. By this Bressers states that the behaviour of the target group is not a sufficient factor in explaining the effectiveness of a policy. This point clearly refers to the distinction between impact effectiveness on one hand and target group and institutional effectiveness on the other hand.

Indirect influence. Because environmental policy relations are not always clear, it is in most cases impossible to find clean-cut relationships between behaviour, instrument and impact. Bressers continues to state that because of the existence of this factor, the government more and more uses forms of indirect influence. This issue clearly refers to the issue of causality.

These elements might influence the effectiveness of the policy instrument and therefore will be given further attention in the evaluation methodology.

3. The concept of causality

When evaluating the effects of a certain policy, one of the central problems that arises invariably is the issue of causality. Indeed, when we say that a certain change in an environmental indicator, or a certain behavioural change is \textit{caused} by policy, what exactly do we mean? To evaluate policy effectiveness, it is not quite sufficient to measure the amount of \textit{objective attainment}. If we’re not able to establish a causal relationship between policy and result, we can’t even talk about policy \textit{effects}; in this case no result whatsoever can be ascribed to the policy with certainty.

According to Coenen (1991), three conditions have to be met before we can establish or demonstrate a causal relationship in an idealtypical fashion. The first one is very logic: the \textit{sequential relationship} between cause and result: the cause has to precede the result. The second condition is the \textit{covariance} of cause and result. This means that they have to be empirically correlated with each other: e.g. more A, means more B. The final condition is the lacking of other explanatory factors. If there is a change in both policy and effect but there are other factors, which can explain the change in effect, then we cannot prove an absolute causal relationship between policy and effect. Only if we manage to eliminate, filter-out, or determine the full interference of all the “other explanatory factors”, it is possible to establish a causal relationship.

Four levels of causal connections are generally distinguished:

1. Necessary and sufficient causes: X is required and will alone lead to Y. Example: if a local authority has the objective to have an environmental official, hiring one will be a necessary and sufficient cause.
2. Necessary but not sufficient causes: X will, under the right circumstances, lead to Y. Example: physical training will lead to an improvement of physical condition, on condition that you don’t get injured or ill and provided you follow a good training program.
3. Sufficient but not necessary cause: X will lead to Y, but so will other things. Example: getting burnt leads to feeling pain, but is not the only possible cause of pain.
4. Contributory Causes: X is neither a necessary nor sufficient cause of Y, but changes the likelihood that Y occurs. Example: levying a CO\textsubscript{2}-tax on gasoline may (or may not) lead to a regression of car traffic. Yet, it is not a necessary or a sufficient cause: other factors can decrease car traffic, and the decrease is not guaranteed either.

For this fourth type, the causal relationship is the weakest, and thus the hardest to establish. Most of the evaluations in the field of environmental policy are dealing with policy issues and circumstances closest to the fourth type. When one tries to establish a causal link between an environmental policy instrument and an
improvement in an environmental indicator (e.g. the amount of nitrate in surface water), it is clear that in most cases the improvement could also be (partially) caused by other explanatory factors, such as other environmental policy instruments, policy in other fields, changes in behaviour of target groups, economic changes, or even completely exogenous factors.

Two types of ‘other explanatory factors’ can be discerned: alternative and competing factors. Alternative factors reinforce the policy effect, while competing factors have an opposite impact on the effect. For example: when evaluating the effects of an environmental tax on the amount of domestic waste, the invention and commercialisation of a new, lighter type of food packaging could be an alternative factor, because it contributes to a decrease of the amount of waste. On the other hand, the demographic evolution towards a society with an rising number of one-person-households, could lead to a sales rise for prepacked meals, which could in its turn end in a growth of the amount of domestic waste. In this case the demographic evolution would be a competing explanatory factor. We will address the issue of other explanatory factors in environmental policy evaluation further in this paper.

4. Specific elements of environmental policy making and effectiveness evaluation

We have defined some key conceptual issues for environmental policy evaluation. We argue here that the substantive field of the environment adds to the difficulties for evaluators, because of a number of characteristics. In fact, environmental problems have some special features, which according to many studies also make them particularly hard to study as policy issues (Mickwitz, 2000). Lafferty and Meadowcraft (1996b) list four features: knowledge deficit, complex geographical pattern of impact and causation, distributions of cost and benefits and time-scale effects. Weale (1992) adds four characteristics in relation to pollution control: it’s a matter of providing public goods, technical complexity, time-scale is often very long and it cuts across sectors if solutions are sought at sources rather than by dealing with effects. Mickwitz (2000) summarises these characteristics in the following list:
- They involve public goods and externalities
- They have been formulated as problems through a process where scientists have played a large role
- They have long time frames
- They are complex
- They have a large amount of different effects that are problematic to commensurate
- They involve huge uncertainties
- They concern geographically remote regions
- They involve stakeholders with conflicting objectives and different belief systems; and
- They have very unequal distributions of impacts on different groups in society

We would like to add two characteristics:
- The irreversibility and/or the existence of thresholds; and
- The fact that the sources of environmental problems are often diffuse

For the purpose of this paper we will only elaborate on these last characteristics. Irreversibility certainly applies to some environmental problems. Degenerated topsoil in, for instance, the desert area of Africa can be characterised as irreversibly damaged: restoration is not feasible in a time period that would fall under even long term policy planning. More absolute is the case of species extinction. Another example, but less absolute, is the pollution of drinking water with nitrates. Once a certain threshold is passed, in casu a concentration of nitrates in drinking water, It will take nature a very long period to restore the natural balances, while the environmental and human consequences are very serious. The depletion of the ozone layer has global consequences but any regeneration, even in optimal conditions will probably take many decades. The irreversibility or at least the very long time frame necessary to remediate,
sometimes very substantial, environmental damage is a difficult dimension in light of effectiveness evaluation. The policy goals, the impact and the societal effectiveness will have to evaluated in light of this element. The evaluation of CO2 reduction policy gets a very different dimension taking into account the fact that scientists (UNPCC) claim that we ought to reduce emissions by 30 to 50 % instead of 5 or 7 %. How effective against this knowledge can the later policy be? Key elements to be included in policy evaluations of issues with these characteristics include the operationalisation of the precautionary principle, the element of time in the policy, and the need for specific monitoring needs.

Pollution, a typical environmental problem can have point or diffuse sources. Point sources are relatively easy to tackle. However, policy actions against diffuse source are a lot more difficult. In water quality management the influx of polluted water from enterprises is easier to control and to treat then the influx coming off households. It is clearly easier to control 10 large discard points then 1000 smaller discard points. Another example is air pollution by dioxins. An enterprise emitting dioxins through a chimney is a point source and relatively easy to control but the emission of dioxins because of traffic is a lot more difficult to control, because of its diffuse nature. The problem with diffuse sources it primarily an identification problem. It is not so difficult to identify point sources, however its is very demanding to identify every small or large source. A second aspect is the cost of control. Controlling a point sources involves a certain cost while will be determined by the influx of pollutants and the transportation costs. The cost of controlling many diffuse sources will involve a higher cost because of the difference in transportation costs.7 In the light of an effectiveness evaluation this characteristics poses some specific problems. Dealing with the diffuse character of some environmental problems is difficult from a pragmatic policy-making point of view as we indicated higher. Evaluating environmental policy, in this case, is also faced with problems. One essential element is the often lacking information needed to evaluate polices aimed at diffuse sources. Diffuse sources mean diffuse or no information. An example can be found in the link between diffuse air pollution and surface water quality. Even in the Netherlands, where environmental information is generally of good quality, there is a serious information gap on this topic. In the best case, the information is retrievable but very expensive. In other instances the evaluation has to be based on scarce and fragmented partial information sources. Since any evaluation of a policy is depending on good data and information this characteristic is clearly a hurdle for evaluating effectiveness.

Measuring environmental policy effectiveness is certainly more difficult the more the issue contains characteristics as listed above. Especially long time frames, complexity, manifold effects, uncertainty and diffuse sources make effectiveness evaluations very complex.

Because of long time frames, the effects of a problem or a measure to counter the problem are hard to establish. An example is the link between the reduction in the emissions of ozone depleting substances and the actual state of the ozone layer. According to the IMO a time lag effect of about 50 years is at stake. This makes the evaluation of the policy rather tricky.

One of the most challenging problems one is faced with in evaluating environmental policy instruments is the impossibility to measure certain effects and forms of effectiveness in light of the characteristics of environmental problems. This can be because the effects have not yet taken place, because the effects are very difficult to measure or because nothing has been done to monitor the policy and its effects. If we do not have an idea about the impact effects, we will have great difficulty giving an appreciating of the environmental effectiveness of the policy instrument. But, not only impact effects are often difficult to measure. Evaluators are also often faced with obstacles measuring some outcome effects. In some cases the target group is the entire population. Measuring the behavioural of a large group of actors can be a peculiar project. Measuring behaviour requires a comprehensive monitoring apparatus or other elaborate methodologies (e.g. longitudinal sensus as in the case of labour market policies). Such methods are scarce.

---

7 To be able to compare the cost, we set the influx of pollutants at the same level for both the point source(s) as the diffuse sources.
in Flanders. In the case of lacking data, the problem of causality is at its strongest. To which degree is the ill-measured change in behaviour attributable to often ill-monitored output effects of the policy?

Yet, output effects are the easiest to measure because of their higher and more direct visibility to the policy makers. In most cases the output effects are evaluated within the framework of an internal administrative evaluation. In such cases the evaluation is part of an internal provision to evaluate the policy and relies mostly on indicators that adhere to the implementation and execution of the policy. One of the problems here is that this type of administrative monitoring and evaluating follows an administrative logic, e.g. are documents sent in, or, are all formal requirements met. This is not always useful for a more comprehensive policy evaluation which aims to establish causal links and places instruments and policies in a broader context.

Evaluating environmental policy within this framework, which is characterised by the above-mentioned problems, is a challenging job. To counter the problems we suggest a number of pragmatic methodological solutions to the problems, which we have discussed above. We will make use of proxies, the design of a causal chain, and in the end an indicator set to provide an answer to the basic question: ‘does the policy work?’. In the next chapter we will further develop these suggestions.

5. Towards a methodology for evaluating environmental policy

In the conceptual part of this paper, we already dealt with the issue of causality. Obviously, the causality problem strongly influences the choice of evaluation methods. The question “is there a causal relationship between the cause and the result?” is not by definition always problematic. For instance, when the government sets up a program for building roads, the causal link between the cause (construction program) and the result (more roads) is obvious. In other cases, as we have indicated in part 4, the link is less clear-cut or often nearly impossible to establish. When this is the case, we have to look for a method, which allows us to solve or at least alleviate these problems.

The examples of this difficult relationship are countless: it is very hard to say whether an improvement of the surface water quality is the effect of a tighter surveillance on illegal dumping by companies, or of other factors, for example better sewage treatment or the abolition of phosphates in detergent.

Depending on the difficulty of the causal relationship, two groups of evaluation methods are available. The first group are the rationalistic methods. These methods are closely related to the classic laboratory experiments. The scientist/evaluator tries to manipulate as many situational variables as possible, thus isolating one particular cause-effect relationship out of many. The second group of methods is called the hermeneutic approach. These methods don’t try to “explain” the causal relationships; the focus is rather on “understanding” the relationships. There is a nuance between explaining and understanding. Explaining involves a higher level of analytical insight.

In literature, the experimental research design is considered to be the most “elegant” to test the link between policy and goal attainment (Coenen, 1991). Unfortunately, a strictly experimental research design is hardly ever practicable as far as policy evaluation is concerned. For instance, from a practical and ethical point of view, it is not possible to abstain a certain policy from part of the population (e.g. no taxes on household waste versus taxes to evaluate the effects of taxes). In most cases, the evaluator will have to use a quasi-experimental or a non-experimental evaluation method. We will further elaborate on non-experimental research methods.

In an optimal setting effectiveness can be established in a pretty straightforward way. One needs to prove the causal link between the different effects and the policy and appreciate the effect in relation to the policy

---

9 Such as in an experimental setting
goal. However the challenge lies in the effectiveness evaluation in situations where the causal link is less straightforward. The EEA (1999b) states that in such a case, and when the causal links are few and predictable, standard models may be derived from examining a small number of case studies. In case the application of a measure is differentiated in space or by target sector, and this can be either within, between or within and between member states, comparative case studies can be used to help identify causal relationships. Another option is the use of in-depth interviews. This technique is especially valuable when the target sector is relatively small. Lastly, and this will be our methodological starting point, when the link between a policy and its impact on the environment is too diffuse or extended and therefore the effectiveness of the impact effects not directly measurable, the focus of the evaluation can shift to immediate outcomes and outputs as a rough proxy for the impact effects. All these methods adhere to the hermeneutic group.

In the next paragraphs we will provide for methodological suggestions to strengthen the last approach. The goal is to go further that the “rough proxy” approach in the case of environmental effectiveness.

**Elements and suggestions for a practical methodology for effectiveness evaluation of environmental policy**

In the following years, the CEEPS will engage in a thorough policy evaluation of concrete Flemish environmental policies on household waste management and water policies. The issues discussed above are the result of a first analysis of the environmental policy evaluation literature. Based on this literature and on a number of policy evaluation projects that have been done by the HIVA research group (Bachus et al, 2001; Gysen et al, 2001a; Gysen et al, 2001b) we will further develop a methodology that allows us to evaluate Flemish policies emphasising the measuring of effectiveness. The suggestions below have to be seen as steps in a process of capacity building for these purposes, not as a final point or as definitive.

The causality between policy and effect, is the tricky part in evaluating the effectiveness of a given policy. The easy position is to throw in the towel. Two positions are possible here. One is an essentialist position based on traditional epistemological reasoning: the proof of causality is impossible in the social sciences. The other position is less deterministic. The argument is that effectiveness evaluation is impossible in this type of complex policy contexts. In stead of accepting these positions, we will try to tackle this problem by researching the policy and its effects in a comprehensive analytical-narrative way. By placing the specific policy (objectives, instruments, effects) within a comprehensive context or policy domain, including actors, the institutional context and instruments we try to come to a ‘reasonable explanation of the causal relationships’. Such explanation is based on ‘reasonable causal chains’ which are contextually constructed (Dessler, 1994).

We will use the framework of figure 1 and try to identify the different elements needed for such an analytical narrative. Identifying the different elements is only the first step. The major evaluation step is to link the impact effects to the policy objectives. This linkage, as we have stated in paragraph 2.1, will provide us with an appreciation of the environmental effectiveness. In most cases this will be very difficult to do, because of the key characteristics of environmental problems and policies. In such a case we will try to show whether there are enough reasons to believe that there is a causal relationship. We will attempt to provide additional reasons through the use of a *causal chain*. This chain is a schematisation of what we call a ‘reasonable explanation based on analytical-causal context analysis’.

It starts by examining and linking the impact effect to the policy objectives. An explanation of the why and how of the objectives (in the policy context) is necessary. Given the desired implicit or explicit impact of the policy we will look for linkages between policy goals and policy impact based on instruments, processes and implementation. If any difficulties arise at the stage of impact determination we will look at proxies for the impact effect. These proxies can be either outcome or output effects. The outcome effects will give us an ‘appreciation’ of the target group effectiveness and –so we hypothesise– an ‘indication’ of the environmental effectiveness. In an identical reasoning, the output effects provide us with an
‘appreciation’ of the institutional effectiveness of the policy. But used as proxies for impact effects, they can also give an ‘indication’ of the environmental effectiveness of the policy. The crucial element in our methodology is the determination and description of the causal strength of the chain.

In other words, the causal chain is a sequence of various effects which can ‘reasonably’ be linked together. We use what Homer-Dixon (1993) called *causal proximity*, as a measure to appreciate the ‘reasonability’ factor. The causal relationship can be anywhere between distant or proximate. Clearly, the description of outputs as proxies will be more distant then outcome effects. The legitimacy of our claims regarding causal links between policy and impact will depend on logic and the empirical evidence of the causal narrative. In addition, given the fact that the narrative is placed in a comprehensive contextual setting, we claim to eliminate in an indirect way as many of the alternative factors as possible.

The more convincing the causal narrative, the more we can talk about a ‘proximate’ causal relation between policy and impact This method is based on the fact that a good evaluation based on these principles makes it increasingly less likely that other contributing factors are indeed playing a large role in explaining the impact of the policy.

**Figure 3**  
Causal relations in causal chain

![Causal relations in causal chain](image)

Figure 3 indicates the causal proximity we can expect between objectives and impact effects, outcome effects or output effects. The direct proof of links between impact and policy will be the most proximate measure or explanation of the effectiveness. Outcome effects and output effects will be more distant than the impact effects in explaining the impact effectiveness of the policy in the light of the objective.

Besides the attempts to establish direct links between *policy effects and the policy*, *other elements* can serve as instruments for effectiveness evaluation. These elements are visualised through the bottom part of the ‘objective’-rectangular in figure 3. The degree of institutionalisation and organisation of the policy, and in particular the instruments and the choice of the instrument mix, but also the implementation and execution processes contain various elements which can serve as indicators. They are part of the causal narrative and can be analysed using various types of what is called plan and process evaluations. Blommestein (1984) indicates that the quality of these policy elements is a good proxy for the effectiveness of the policy. However, he warns for the danger that lies in the other explanatory factors and the causal strength (or weakness). In applying our methodology we try to account for this concern, especially with the choice of the proxies and by using causal narratives and chains.

---

10 This list is not exhaustive
In the end we hope to be left with a very small portion of the policy which can not be explained by the policy effects or by the other elements. Clearly the smaller this part is, the closer to a causal explanation we have come.

6. Conclusion

Measuring the effectiveness of a policy instrument in a complex context such as environmental policy is no sinecure. Effectiveness of environmental policy is not be mistaken for institutional effectiveness, which is only one of four kinds of effectiveness. The bottom-line effectiveness evaluation in environmental policy is impact effectiveness. Impact effectiveness compares the impact of the policy to the state of the policy field under review with the objectives of the policy. This is fundamentally different from target group, institutional and societal effectiveness.

Measuring this ‘impact effectiveness’ is a challenge given the fundamental problem of causality. In other policy fields, causality creates obstacles of effective evaluation, but in this paper we claimed that environmental policy suffers especially from the influence of a number of inherent features. Taking these obstacles into account we have provided some elements for a methodology based on the hermeneutic method. The basic idea of our methodology lies in the attempt to provide a ‘reasonable story’ or an analytical-causal narrative. Outcomes and output effects are used as proxies for impacts and are placed in a policy context.

To conclude and to prove our point of the importance of a methodology to measure impact effectiveness, we would like to quote Per Mickwitz (2000): ‘If evaluations of environmental policy instruments are undertaken without due consideration of the specifics of environmental problems there is a significant risk that little impacts and low effectiveness will be found’.

We hope that through the period of the next several years, we will be able to establish a number of criteria and indicators that are useful for effectiveness evaluation. In line with the reasoning that is followed by the EEA, we wish to tackle the issue of effectiveness evaluation in the environmental policy field in a more systematic way. This goes against the argument of some policy analysts in the field who claim the impossibility of this exercise. However, given the importance of environmental policy making and given the inherent and also political legitimacy of the question ‘does this policy work’ we wish to at least come to better methods.

7. Bibliography


Deelen, F. van der, (1993), *De gereedschapskist van de overheid: een inventarisatie* in: Beleidsinstrumenten bestuurskundig beschouwd, Assen/Maastricht.

EEA, (1999a), *The contribution of EU monitoring and reporting to ‘sound and effective’ environmental policy-making*. First draft, part of the REM project, Copenhagen.


