

Policy Evaluation and Environmental Policy

The phenomenon of policy evaluation raises various questions with environmental policy-makers and with those civil servants and consultants who advise them: What is policy evaluation? How can it help (me)? How should policy evaluation be conducted? Which methods are available? The aim of this handbook is to help answer such questions in a practical way.

Anticipating the systematic presentation of a series of methods for (environmental) policy evaluation in Chapters 3 and 4, we first need to consider and to elucidate some of the basic concepts involved in policy, in policy evaluation and in the particularities of environmental policy evaluation. This chapter introduces some basic and yet fundamental concepts of policy evaluation, whereas Chapter 2 deals with the particular features of environmental policies and the specific issues that may arise from them.

The concepts and the elucidations dealt with hereafter are directly related to the actual definition of policy evaluation that is applied in this handbook: *policy evaluation is a scientific analysis of a certain policy area, the policies of which are assessed for certain criteria, and on the basis of which recommendations are formulated.*

This definition contains certain elements that are crucial to the reader's understanding of this and subsequent chapters: policy, analysis, evaluation, criteria and recommendations:

- *Policy*: while everyone has some concept of what policy is, it is worthwhile considering some frequently applied definitions, which we will do in this chapter. Depending upon how 'policy' is viewed, a different perspective emerges on the notion of policy evaluation (see Section 1.1 'What is policy?').
- *Analysis and evaluation*: in order to evaluate policy, one must first analyse it. This entails research into the what, how and why of a specific policy context. Such analytical questions always precede the actual evaluation: how effective, how fair, how enforceable and so on, is a given policy?

- *Criteria*: the essential difference between policy analysis and policy evaluation is that the latter is based explicitly on a set of specific assessment criteria. The precise nature of these criteria needs to be determined anew for each individual setting. The evolution of policy evaluation reflects a gradual change in the type of criteria applied (see Section 1.2 ‘Criteria for policy evaluation’).
- *Recommendations*: policy evaluation is generally not an end in itself. Its purpose is, rather, to improve policy in one way or another, even when the impact of evaluation studies on actual policies is an issue in itself, which will be briefly touched upon.

Finally, we need to point out that our definition of policy evaluation implies that three types of knowledge come into play when evaluating policy: analytical knowledge, which describes and explains (what is ...?; how does ... work?; why is ...?); evaluative knowledge, which assesses (how good is ...?; how suitable is ...?); and prescriptive knowledge, which recommends alternatives (how might ... be approached differently?). This handbook focuses deliberately on methods designed to generate evaluative knowledge about policy processes, policy products and policy fields. It devotes no or very little attention to the technicalities of methods. In other words, it leaves aside questions such as how do I conduct interviews? How do I reconstruct a policy process? Nor does it touch upon methods of policy prescription, with questions such as how should I formulate recommendations? How should I test the feasibility of these recommendations? What the reader will find, though, is a broad outline of various policy evaluation methods in Chapters 3 and 4.

This chapter consists of two sections. First, light is shed on some of the key concepts: ‘policy’ and the different views on policy are the focus of Section 1.1; ‘policy evaluation’ and the various criteria involved are discussed in Section 1.2. In order to prevent the chapter from being a purely theoretical survey, we provide boxes with various well-known and less well-known definitions; give concrete examples; and point out practical problems. For the same reason, we use very few references in the chapter itself, but we provide an overview of some of the publications we have used as our main sources of inspiration at the end of the chapter. More detailed references to specific scientific literature are provided in Chapters 3 and 4.

1.1 WHAT IS POLICY?

1.1.1 The policy cycle

From the huge number of books in policy studies and policy analysis, one can draw an even greater number of definitions and concepts of policy. All of them attempt to grasp the very heart of the complexity that ‘policy’ represents, and to reduce its complexity by using relatively accessible and understandable models.

One of the most popular concepts in policy studies is to simplify the policy-making processes into a series of stages. This conception originates from the early days of policy analysis, and has been altered and sophisticated since, but essentially draws on a parallel with production processes. It conceives policy-making as an ongoing iteration and reiteration of a policy cycle (see Figure 1.1), in which more or less consecutive stages and the according policy processes can (analytically) be distinguished.

To further elaborate upon the policy cycle idea, we briefly discuss some of its steps:

- A number of *problems* or *societal focal points*, experienced to different degrees by people or groups, are highlighted socially and politically. In a number of cases, this may lead to the problem being placed on the *political agenda*. This is the *agenda-setting* phase. As the political agenda constantly tends to be overloaded, the existing problems are *selected* and *prioritized*. This implies that certain problems may be removed from the agenda, while others may be reformulated.

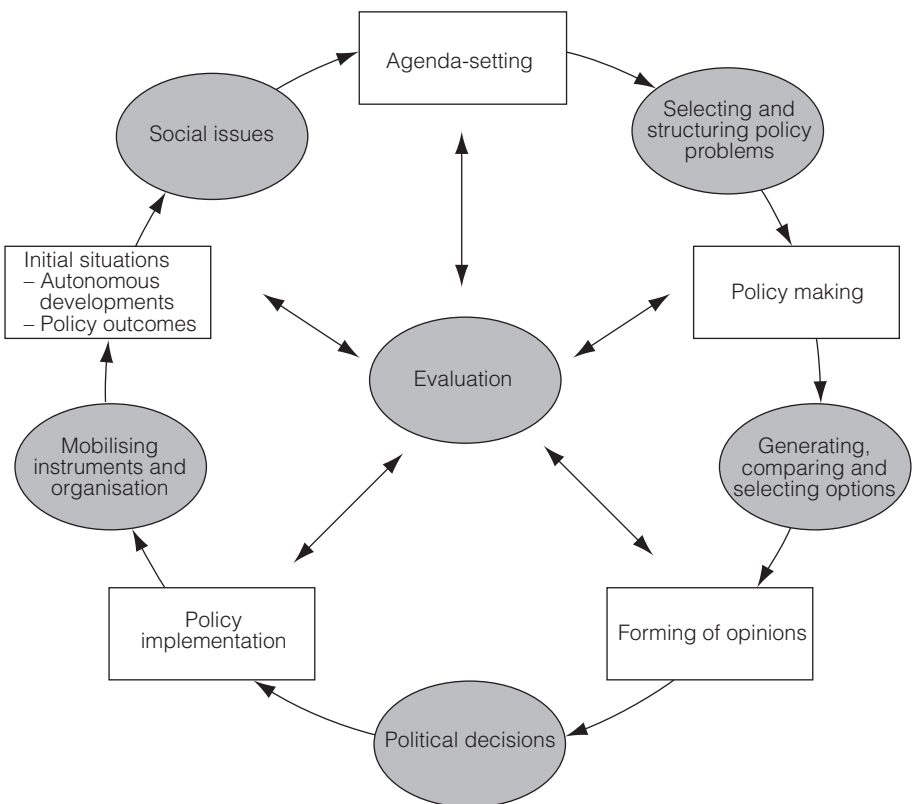


Figure 1.1 *The policy cycle*

These processes of problem definition and selection are determined by (environmental) incidents, by deliberate agenda-setting and problem definition strategies, and by the structural characteristics of the political system, which make it more sensitive to some problems than to others.

- In the *policy-making* process, solution options are developed. These solution options are presented, selected and compared in the media by interest groups, citizens and civil servants, as well as within political circles. Sometimes, scientific considerations or methods, such as cost-benefit analysis, also come into play. Opinions start to form, advocates and opponents begin to group together, and political pressure in favour of divergent solutions gradually builds up. In short, various *policy options are generated and selected*. 'Good' problem analysis and 'good' strategy formulation must, after all, promote a successful approach and its implementation. As a matter of fact, problem analysis and strategy formulation run largely parallel to one another.
- *The forming of opinions* on the remaining options takes on a more decided character. The main social and political forces adopt a position on the various solutions. This is not a 'neutral' choice, but an entirely political one: although goal-oriented, it is partly informed by political and societal viewpoints (on the role of government, the responsibilities of the business community, etc.), and, of course, by the balance of power between various social and political forces in society. This forming of opinions leads to *political decisions*, in which the basic ideas of the policy, its aims and its strategies are set out.
- The next phase is that of *policy implementation*. Basic political decisions must first be 'concretized' into specific measures: mobilizing and setting aside the required resources (manpower and budgets), specifying procedures and rules, and determining the division of tasks and the coordination between the various organizations entrusted with implementation. In sum, it involves *a mobilization of instruments and organizational planning*, the actual functioning of which is also referred to as 'policy output'. The question remains, though, whether all the resources, knowledge, power, legitimacy, etc. that are necessary for implementing the policy are, indeed, available. Furthermore, policy implementation is partly down to actors (government organizations, companies and others) who are not necessarily in favour of policy action, which entails a risk that implementation will be hindered or distorted. Policy-making is a multi-actor and a multilevel phenomenon, which partly explains the sometimes substantial differences between policy on paper and policy in practice. While the latter may be referred to as 'policy performance', the differences between ambitions and actual performance are known as the implementation deficit. This may, but need not, have a detrimental effect on the success of a policy.
- In the fifth stage, the policy *effects* manifest themselves. First and foremost, effects concern the intended changes in the behaviour of the target group (such as companies, households and motorists). This is referred to as the 'policy

outcome'. Second, the term effects may refer to the intended environmental impacts (i.e. to an improvement in environmental quality). For that matter, the relationship between the two is not easy to identify. After all, a number of the outcome effects are only partly realized or not realized at all. And the end effects, the effects on the environment, are often realized only in the long run. Chapter 2 will go into these difficulties of evaluation that are partly specific to the environmental domain. In addition to these intended effects, unintended side effects may be generated that can potentially neutralize the intended effects. The controversial role of biofuels in recent debates and policies on climate change, including their unintended effects on agricultural policies, and therefore on global poverty and development, provides an example.

- In addition, while the policy process goes on, simultaneous and rather *autonomous societal developments* are occurring as well: while the environmental policy is being implemented, economic growth, individualization, globalization and other societal developments continue to unfold. Simultaneously decisions are taken in other policy fields, such as spatial planning, traffic and others, that will have an impact upon either the environmental quality or upon the target groups' behaviour. These developments, together with the outcomes of the environmental policy process itself, will lead to newly emerging social problems and therefore to renewed problem formulation and agenda-setting. In short, the policy cycle will be reiterated.

BOX 1.1 TYPES OF POLICY EFFECTS

The question of whether (environmental) policy is having an impact – many more aspects of which will be discussed below – inevitably gives rise to a counter-question: what kind of impact?

Traditionally, a distinction is made in the literature between:

- *Policy output or policy performance*: this entails the quantity and quality of the products and services delivered by policy-makers – for example, the number of permits granted, the revenue generated through environmental taxation, actual expenditures on environmental subsidies, the amount of information provided, etc.
- *Policy outcome or social change (also referred to as behavioural effect)*: this comprises policy impact in terms of behavioural change among citizens, companies, farmers, consumers and other groups in society – for example, modified driving behaviour, private investment in water treatment infrastructure, increased purchases of biological products, etc.
- *Environmental impact or environmental change (also referred to as ecological impact)*: examples of this include a reduction in nitrogen oxide (NO_x) or sulphur dioxide (SO₂) emissions, a decrease in water pollution or more land for organic farming.

The policy cycle is a useful concept and metaphor. After all, it reveals the different ideal-typical stages in the policy process and, in so doing, the policy-making activities that theoretically take place in each of those stages. The policy process is dissected into a kind of production process involving various sub-activities, each of which requires its own raw materials and treatment methods, as well as its own semi-finished products. This means that the activities in each of these policy stages can be methodically analysed and substantiated. What we are dealing with, then, are methods for policy development *and* for policy analysis. Many books are available in which, either for the entire policy cycle or for parts of it, various methods are described that may be used to support policy development. At the end of this chapter, a list is provided of some overview publications, while Chapters 3 and 4 refer to more specialized methodological literature.

BOX 1.2 NEITHER *EX ANTE*, NOR *EX POST*, BUT RATHER *EX NUNC*

A traditional distinction in policy evaluation is that between *ex ante* and *ex post* evaluation. The concepts are self-explanatory: *ex post* evaluation refers to the evaluation of policy after it has been developed and implemented, while *ex ante* evaluation means that policy is evaluated prior to implementation (i.e. while it is still very much on the drawing board).

Three comments are in place with regard to this distinction:

- 1 Methods for *ex ante* evaluation are, in fact, at once methods for policy design and policy development. Many handbooks deal with them under a single heading.
- 2 It is clear that *ex post* policy evaluation leads to new insights for policy development and *ex ante* policy evaluation. This learning effect, however, requires a minimum amount of *ex post* policy evaluation research that should, furthermore, provide sufficient generic (and not simply case-specific) insights. In Chapters 3 and 4 we also discuss approaches to policy evaluation in which specific attention is paid to this learning effect.
- 3 Most policy evaluation research is neither *ex ante*, nor *ex post*, but rather, *ex nunc* or *ex in between*, if you will: it deals with the evaluation of current policy, the interim evaluation of which may lead to modifications. Since policy is rarely completed, let alone terminated, there is very little *ex post* policy evaluation in the very strict sense of the word. Still, we will persist with the distinction between *ex post* and *ex ante*.

In the context of this book, a further aspect of the policy cycle and the position and role of evaluation is important. Many textbooks look at 'evaluation' as the penultimate stage of the policy cycle, with its final outcomes and the impact of

simultaneous societal developments conceived to be the last stage. Other textbooks and our Figure 1.1, however, place 'evaluation' at the centre of the cycle, suggesting that each phase of the policy cycle may be evaluated. The methods we list in Chapters 3 and 4 differ in many respects, including that some of them are appropriate to one or, in particular, several of the policy cycles' evaluation. It might be clear, indeed, that policy evaluation questions may relate to each or to several of these phases. By way of illustration, we consider some environmental policy problems with examples of evaluative questions about successive stages of the policy cycle:

- *Problem formulation: livestock manure.* From the late 1970s until today, different European countries have faced environmental problems related to intensive livestock farming (i.e. problems of increasing volumes of manure). From international comparative research on this period, it is clear that these problems have been defined successively as a non-issue, a regional surplus, a processing problem and a matter of volume control. For domestic reasons, varying from a country's physical geography to the power relations between its main stakeholders, the problem has further been defined in terms of water pollution, acidification, an odour problem, a technical problem within the agricultural sector, etc.

From a policy evaluation point of view, this raises interesting questions as to how the prevailing problem formulation has affected the subsequent phases of the policy development process and, ultimately, the effectiveness of this policy. In other words, a legitimate evaluative question could be: to what extent was the prevailing formulation of the livestock manure problem adequate (appropriate, consistent, complete, etc.) in order to pursue an effective policy?

- *Problem formulation/selection: the role of public perception.* Under normal conditions, environmental policies and their risk assessment are assumed to be formulated by experts. They tend to present risks in terms of probabilities and impacts. Under situations of environmental crisis, though, public perception, echoed by the media and parts of the political system, tend to have a huge societal and political impact. Risks related to the environment, food safety and public health, although assessed to be small by experts, may then be magnified as examples from the BSE crisis ('mad cow' disease), foot-and-mouth disease and others illustrate.

From an evaluative point of view, this raises questions on the respective roles of social perception and scientific information in assessing an environmental problem, its formulation, its priority on the agenda, and its impact on the actual policy measures taken.

- *Instrumentation: wastewater policies.* Almost all countries on the globe face more or less huge wastewater problems. The way in which they deal with it is quite different, of course, in terms of their geographical, environmental,

economic and political situation, as well as their capacities. However, these political responses will also differ with regard to their basic strategies: do these responses predominantly rely on regulatory strategies, economic instruments, or public or private investments and efforts, etc.? The strategies adopted are partly inspired by considerations of effectiveness and efficiency; but no doubt these have also been determined by domestic institutional conditions, such as the role and position of the state and its different administrative levels, as well as the role of water companies, private investors, etc.

In terms of policy evaluation, an intriguing series of possible questions rises: to what extent has the choice of instrumentation strategies been inspired by institutional needs and by the pre-existing organization of the policy domain, rather than by considerations related to effectiveness? In other words, is the instrumentation of the policy that was opted for the most appropriate (effective, suitable, legitimate, accepted, etc.)?

These (real life) examples highlight just a few of the policy evaluation questions that may be formulated in each stage of environmental policy-making processes. Chapters 3 and 4 contain a variety of methods, some of which specifically focus on a particular stage of the policy cycle, in some cases also with specific methods. As might be clear from these examples, however, many of the evaluative questions can only be answered adequately if one analyses and assesses a much larger part of the policy cycle.

While this demonstrates that, as a concept, the policy cycle definitely has its merits, its simplicity can also be misleading. More specifically, it may create the impression of not being an ideal-typical representation, but rather an accurate reflection, or even a normative standard, of reality. As far as we are concerned (and in the context of this book), the policy cycle is no more than a heuristic or analytical tool. In no way do we wish to imply that processes of policy-making actually unfold according to these (chrono)logical steps, nor, indeed, that they ought to. There is sufficient empirical evidence from policy studies to falsify any of these suggestions. Furthermore, the notion of the policy cycle, like any schematic representation, tends to accentuate some, and to ignore other, parts and characteristics of policy reality. The following section considers varying answers as to what policy is, or at least how it could be conceived; from there, it relates divergent views to (environmental) policy evaluation.

1.1.2 Policy: Three different views, three types of evaluation

The questions of what policy is and how it can be approached have given rise to much debate in policy studies. Consequently, many scientific publications on the subject carry titles such as 'Views on policy' or 'Public policies: Competing

paradigms'. Even in everyday language, various definitions of policy are used, which again are based on different perceptions and views. We do not intend to provide an extensive overview of these discussions in policy theory and practice, but merely to discuss three clearly distinguishable views on policy, each of which implies a different perspective on policy evaluation and its functionality. The views that we discuss are 'policy as a control loop', 'policy as political interaction' and 'policy as an institutional phenomenon'. We provide a brief outline of each perspective and its implications in terms of policy evaluation.

Policy as a goal-oriented rational-synoptic process

In the first conception, policy is seen as a process of conscious and rational problem-solving. In policy cycle terms: a social problem has become a policy problem as a result of a process of agenda-setting, with a clear problem definition and consequent goal-setting, on the basis of which different policy options are developed. These options are compared and weighed against each other in some sort of multi-criteria evaluation, after which one or a combination of them is selected and implemented. At the end of the process, a policy evaluation can assess whether, or to what extent, the initial problem has been resolved, and whether or not the policy cycle must be reiterated.

In this view, policy is like a control loop. It is a concept borrowed from the simple principles of process technology and engineering sciences, the household thermostat being a well-known example. When set at 20°C, the thermostat will detect 'a problem' at 19°C (i.e. a difference between the actual and desired situation). This elicits a signal: 'heat on'. Consequently, heat is produced and transported through the system, heat exchange ensues and the room temperature rises. Once the desired room temperature has been reached, 'heat off' is signalled and the problem has been solved. Should the problem reappear, the process will be repeated. In short, the control loop resembles a fully automatic operational process, with feedback of information resulting in task modification.

No policy scientist takes the control loop metaphor too literally. There are, however, many policy scientists – and even more policy practitioners – who tend to regard policy-making as a process that is driven mainly by a problem-solving rationale. In this view, policy is approached as a means of rationally tackling social problems. The word 'rational' refers, first and foremost, to the method: effective and efficient strategies are sought on the basis of scientific methods. Multi-criteria analysis and other approaches of *ex ante* policy evaluation are ascribed a major, if not decisive, role in these processes. 'Rational', secondly, refers to the motives and the nature of the parties involved: these parties are rational beings, who – like *Homo economicus* – calculatingly look for the best possible solutions. In short, this concept of policy is strongly influenced by engineering and by economic thought, and elaborated upon in rational choice theories in different social science disciplines.

Seen from this perspective, the purpose of policy evaluation is quite clear: policy must essentially be assessed and optimized using the criteria of goal attainment and suitability, often played out in a series of clear-cut indicators. Conversely, a great many practical evaluation questions that policy-makers put forward are, often rather implicitly, aligned with this engineering-inspired view of policy-making, restricting any assessment and evaluation of their work to the monitoring of a set of indicators. They are, after all, concerned with identifying the best (i.e. most rational) way of analysing, tackling and solving the policy problem. In other words, policy evaluation is intended to contribute to the further rationalization of policy.

Policy science has a large number of methods at its disposal to help answer the kind of policy evaluation questions originating from this perspective. Let us consider some examples, relating to three different policy phases:

- 1 *Reconstruction of policy theory.* Policy is invariably based on certain assumptions regarding the causes of the problem to be addressed, the solution strategies to be pursued and the underlying ethics and moral principles. These assumptions are largely implicit: if they are made somewhat more explicit in policy documents, then this is usually not equivocal. The European Water Framework Directive gives a recent example of a policy document that explicitly lists a series of principles – although very differently elaborated upon and ‘operationalized’, and without addressing the possible and actual contradictions between them. As a consequence of the often implicit character of these policy principles, the evaluator needs to reconstruct them. Such reconstruction of underlying assumptions generally reveals inconsistencies (‘irrationalities’) in analysing the problem, the solution strategies and the premises, and identifies ways of rectifying these shortcomings. Drawing up so-called target trees or goal mean trees are commonly applied tools to check policy assumptions for their validity and reliability. The aim of this type of policy evaluation is to upgrade the quality of policy assumptions to a scientific (i.e. to a more rational) level.
- 2 *Predicting effect and effectiveness.* A rational approach to a policy process implies a prior (*ex ante*) estimation of the likelihood that the intended effects will actually be attained. Using a wide range of methods of forecasting, impact assessment and the like, an attempt is made to predict the effects and, consequently, the effectiveness of a policy. Crucial in this respect is the assessment of the likelihood that the target group (e.g. individual motorists, a group of local communities or companies) will change its behaviour, and an appraisal of the degree of regulation, financial stimulation or public information required in order for such a change to take place. Depending upon the field of study (economics, public administration, psychology, etc.), the approaches differ; but their common aim is to establish which type of tool should preferably be deployed and which ‘dose’ should be administered to achieve the desired

'effect'. This approach, too, is based on a rational view of policy-making as one assumes people and organizations to be sensitive to a change in the balance of wins and losses, and to adapt their behaviour accordingly.

- 3 *Gauging effectiveness.* In this perspective on policy-making, problem-solving is the driving mechanism and the ultimate, sometimes exclusive, criterion for assessment. Thus, *ex post* effectiveness evaluations are clearly the polar opposite of *ex ante* evaluations. A whole repertory of methods is available for answering the effectiveness question, especially from a cost-benefit and, even more so, a cost-effectiveness point of view.

From a rational perspective on policy, two issues are extremely important here: the problem of the causal connection between policy efforts (the 'dose') and their (different types of) effects, on the one hand (see Box 1.1 on the different types of policy effects; see also Box 1.3 on causality), and the problem of side effects, on the other. Here too, various methods are available, both quantitative and qualitative in nature. However, given its premise of rationality, this approach will usually show a preference for quantitative methods.

The perspective whereby policy is seen as a control loop is tempting, particularly its simple basic assumptions and transparency: people observe a problem and decide to tackle it by means of a well-considered policy intervention. Consequently, it should be easy to determine whether that intervention has been successful or needs to be repeated or modified.

At this point, however, it might be clear that this rational view of policy and the metaphor of the control loop entail certain limitations, resulting from their specific conception and reduction of reality. First and foremost, the premises of the 'rational actor theory' are only tenable to a limited degree. The actions of human beings and organizations may be (and actually are) guided by considerations other than those that are rational and goal-oriented. Second, the notion that policy is driven mainly by the desire to solve problems is an idealistic distortion: a whole array of driving forces may come into play in problem formulation, agenda-setting and prioritization, as well as in the implementation of decisions. These range from moral indignation to emotions, from power imbalances to monopoly situations, from private interest to prestige, from policy tradition to disputes over competencies. All of these considerations, motives and contextual factors, regarded as 'irrational' from a strictly rational point of view, do play their role in policy-making, in general, and in environmental policies. This, in part, explains why official policy goals almost never coincide with real, but often difficult to detect, ambitions. Third, by approaching policy-making as a control loop, one effectively ignores part of the reality in the analysis. The one-sided emphasis on the solution-focused nature of policy-making overlooks other aspects, one of which is the fact that policy is not a fully automatic process or control loop, but is, rather, created in an environment of continuous social and political interaction. This brings us to the second perspective on policy-making.

BOX 1.3 THE QUESTION OF CAUSALITY

A political and methodological focal point in policy evaluation, where the evaluator tries to gain insight into the degree of effectiveness, is the main focus of causality. By this we mean that one strives to (positively) answer the question of whether the intended policy effects (the goals) have actually been achieved thanks to the policy.

This sounds much easier than it is. After all, a policy may attain its goals due to favourable external circumstances, or due to the impact of another policy, or because of a foreign influence. A reduction in polluting emissions may be the result of a flagging economy rather than an effect of environmental policy. Policy-makers, in general, tend to attribute policy failure to situations abroad or to 'external circumstances', while they tend to claim responsibility for policy successes themselves.

The difference between goal attainment and effectiveness lies precisely in the causal role of policy. Thus, it is imperative that an evaluation of policy effectiveness is conducted with care and accuracy. However, this is complicated in various ways:

- Policy aims are often unclear, sometimes deliberately vague and difficult to measure so that it is impossible to ascertain unequivocally that the goals have been attained.
- Policy rarely consists of one measure – a so-called single shot – but usually encompasses a whole series of measures, making it difficult to assess the effect and effectiveness of the policy.
- The policy path, from policy formulation to policy outcome, passes along many links and through numerous actors. Chapter 2 discusses the multi-actor and multilevel features of contemporary policy-making. Moreover, especially in the case of environmental policy, there can be a substantial time lag, another specific feature that Chapter 2 addresses. Groundwater quality and halting biodiversity loss are self-evident domains that illustrate the huge time delay between policy efforts and their eventual effects. Consequently, causal chains are long so that it becomes tricky to identify crucial factors.

For all of these reasons, 'pure' examples of effectiveness evaluation are scarce. This explains why one particular example always stands out, including at the European level (EEA, 2001): the effectiveness of initially phasing out and subsequently banning leaded petrol is entirely quantifiable, down to the end effect of 'amount of lead in verges'. It is an exceptional example because it concerns a single substance, a single sector of industry, a limited number of measures taken over a limited period of time, and a fairly simple, natural pattern of diffusion. The effectiveness in terms of public health impact, however, is much trickier to calculate because it involves a much more complex set of influences and outcomes.

Policy-makers find it hard to accept that the question of causality can only be answered in exceptional cases. Consequently, research into policy evaluation is constantly searching for methods to describe the effects and the effectiveness of policy, despite all of the pitfalls such an undertaking inevitably involves. Chapters 3 and 4 illustrate the divergent responses that literature has given to this challenge. In fact, one invariably tries to make plausible what the effects and the effectiveness of policy are (or have been) through indirect reasoning and argument. The latter comes close to what is called 'circumstantial evidence' before a court of law.

Another aspect of policy-making that is neglected is the fact that policy-making unfolds in a pre-set institutional context. This latter point brings us to the third perspective on policy making.

Policy as political interaction

Whereas the previously discussed view on policy is strongly inspired by engineering and economic thought, and by a rational actor model, the 'policy as interaction' perspective we discuss here is more inspired by paradigms from political sciences, in which policy is mainly the product of power relations between various social and political actors, groups, convictions and interests. In other words, whereas in the first view, policy is regarded as a matter of rationally designing and implementing a problem-solving strategy, in the second, it is seen mainly as an unfolding struggle between conflicting interests and power bases.

This view on policy, too, can be elucidated by passing through the successive links in the policy cycle. In doing so, the differences will become clear, not only in relation to the policy view, but also in terms of questions about, and expectations of, policy evaluation:

- *Agenda-setting:* if policy is seen as political interaction, the questions of whether and how social problems can be placed on the agenda depend mainly upon the capability of the various social and political actors to push through 'their' issues and problem definitions. This capability, in turn, depends partly upon the actors' own resources in terms of expertise, personnel, financial and other means, and partly upon the political context in which the policy unfolds. Thus, the previously mentioned successive shifts in the formulation of the livestock manure problems in different countries may reflect the changing balance of power between the actors involved in a context in which, in many countries, there used to be a privileged relationship between the ministry and the farming industry. While in some countries the manure issue has led to the breakthrough of this neo-corporatist policy-making style and practice, other countries witnessed the successful survival of a rather closed agricultural–environmental policy domain. Empirical research in different countries illustrates the relation between the balance of power and the interactions between the stakeholders involved, on the one hand, and the way in which the problem is formulated and handled in the consecutive stages of the policy cycle, on the other.

In order to analyse and evaluate policies from this perspective, one could rely on different models of agenda-setting and issue-handling – for instance, the barrier model – suggesting that successful agenda-setting implies conquering a number of barriers which selectively block access to the political agenda, or the policy streams approach, suggesting that agenda-setting depends upon bringing together problem definitions and solution options, and upon seizing windows of opportunity. These and other approaches

emphasize the capacities of political agencies to mobilize resources and to seize opportunities for the sake of agenda-setting, or they point to the institutional context that offers such opportunities or, rather, blocks these. Any of these approaches can provide insight into who has succeeded in imposing a problem definition and, as the case may be, into associated solution strategies, which power tools have been decisive in this regard, etc. This perspective also provides insight into why, in certain cases, issues are denied or removed from the political agenda, and why no policy response takes shape even if some recognize there to be a clear social issue. The latter might be related, for instance, to the respective impact of experts' assessments versus lay people's evaluations, particularly in some crisis circumstances, which we referred to earlier. In brief, it is not its problem-solving orientation or its 'rationality', but, rather, the way in which the problem formulation reflects political and societal interaction that is at stake in policy evaluation.

- *Policy formation:* if one regards policy as a process of political interaction, the crucial question in this phase is not whether and how the selection of policy options and, ultimately, of a policy strategy, can be reasoned and optimized on the basis of the problem solution. The question is, rather, how this selection process should involve negotiations between the various actors concerned and, conversely, how the approach chosen should reflect the problem perception and solution paths of these actors, as well as the relationships between them. From this perspective, again, the policy goal is not so much the result of an ambition to resolve a problem, but rather a reflection of the power balance between the actors involved. These actors may be political, economic and social interest groups, as well as government agencies, which must themselves promote their viewpoints and fight for authority, resources and influence. If policy comes about in such a multi-actor setting, where different perspectives and interests must vie for influence and power, the question arises how such a tangle of interaction and intrigue may be woven into a compromise.

As far as policy analysis and policy evaluation are concerned, approaches such as network analysis, power relations and power-balance analysis, stakeholder mapping, and the like offer useful frameworks for explaining and assessing policy choices, including understanding the success or failure of these choices. According to this view, policy evaluation is no longer limited to comparing goals with achieved effects. It is, in other words, no longer seen merely as an assessment of policy products (output). This doesn't mean, however, that policy evaluation from this perspective need not be restricted to an evaluation of the policy process, particularly the interactions between those involved. Instead, policy evaluation from this perspective should be able to scrutinize the connection between the policy process and the policy organization, on the one hand, and its very outcome – the substantial policy product – on the other. Much attention is devoted to the manner in which

the actors, either under the guidance of government or through other facilitators, mediators or policy brokers, organize the policy process, and whether the chosen approach is suitable for connecting contradictory interests. In other words, one explores which policy organization and processes are suitable or not for the so-called interweaving of goals. In addition, the evaluation may focus on the acceptance and workability of the results achieved. After all, there are examples of unworkable as well as extremely successful compromises. From a policy development point of view, it can be interesting to investigate *ex post* which circumstances of policy process organization and implementation play a role in this respect.

- *Policy implementation*: according to the rational view on policy-making, the implementation of policy is conceived as a politically neutral process. This results in policy evaluations that investigate where, how and why implementation stalls, and why this was not anticipated more effectively.

If one regards policy as a form of political interaction, however, policy implementation is 'the continuation of decision-making by other means'. Seen from this perspective, policy implementation is a process in which a variety of decisions need to be constantly made, and throughout which power, acceptance, information, capacity and other elements play a decisive role. The policy implementers, who from a rational point of view are seen too readily as fully automated signal relayers, are now viewed as actors with their own goals and ambitions, who are able to carry out or oppose policy decisions from elsewhere, who may or may not possess the resources to do so, etc. Thus, policy can be evaluated in terms of the degree to which it possesses the means of facilitating or obstructing implementation (or, as the case may be, having it obstructed).

In this respect, policy studies and policy evaluation literature recently paid a lot of attention to implementation processes that ask for the interplay between government bodies at different administrative levels: the formal and material implementation of international agreements (e.g. on climate change and biodiversity) is one example; the interaction between national and local authorities reveals similar mechanisms of implementation as a multilevel game. This conception of policies as a multilevel game implies that, again, interaction between a multitude of actors is seen as the key factor in explaining successes and failures of policy processes.

It should be clear from the above that while the perception of policy as a form of political interaction is certainly a more realistic representation of policy-making, it does not make policy evaluation any easier. The increased realism lies in the following aspects:

- There is now room in the policy evaluation for dealing with other, entirely different, motives than those focused on in the rational model. This results in a less one-sided and more complete picture of reality.

BOX 1.4 POLICY: GOAL-ORIENTED OR GOAL-SEEKING? SOME CONSEQUENCES FOR EVALUATION

From a rational perspective, one would assume that policy is goal-oriented. This implies that there is a clear and unambiguous problem definition and solution strategy – which, in reality, is seldom the case – on the basis of which the policy effects can be gauged. In other words, according to this viewpoint, the policy itself provides the criteria against which it is to be assessed.

The fact that policy is regarded as the result of interaction not only means that more attention is paid to the process, but also that policy is seen to be goal-seeking rather than goal-oriented. The latter is especially the case in those policy processes that use interaction as an essential means of getting policy processes moving. This is the case in many, often regional, processes under mobilizing but rather vague umbrellas, such as ‘sustainability’, in which government bodies, market players and all kinds of societal organizations interactively aim at designing and implementing an appropriate policy for a specific region or for resolving a series of interrelated issues.

When evaluating this type of policy, process evaluation must, of course, play an important role. Traditional methods of gauging effectiveness are largely useless in this respect since there is no clear-cut goal definition – or, when there is one, it may shift over time as interactions go on. A policy evaluation of these processes is likely to be more successful when it is designed interactively and is participatory. Such an evaluation can also involve feedback, not only with regard to the actors’ satisfaction with the process, but also whether the actors feel their ambitions and goals are reflected in the end result.

Chapters 3 and 4 discuss forms of participatory evaluation. As we will see, newer variants of environmental policy-making (interactive, participatory, deliberative, etc.) have greatly contributed to the development of this approach to policy evaluation. In turn, participatory designs for policy evaluation might help, in these circumstances, to overcome the limitations of ‘rational’ designs.

- There is also room for evaluating the policy process, particularly the way in which it enables and structures interactions between those involved, as an important explanatory variable for goal attainment and the suitability of policy.

As a consequence, however, policy evaluation seems more complex:

- From this perspective, policy evaluation in terms of goal attainment and effectiveness is no longer sufficient. In addition to these criteria, we must now also deal with conditions such as acceptability, legitimacy and participation, which are primarily intended to help assess the process. This implies a complication: policy may be rational; but it is so rarely accepted that it is not implemented and, consequently, remains ineffective. Conversely, weak

policy processes may result in effective policy. In other words, criteria for policy evaluation may yield contrary results. Section '1.2 Criteria for policy evaluation' deals in more detail with the issue of criteria and the contradictions between them.

- In many cases, particularly if there is lengthy interaction and cumbersome political decision-making, policy evaluation in terms of goal attainment or efficiency is virtually out of the question. Such a situation tends to arise when policy goals are vague, ill formulated or frequently reformulated. Lasting interactive processes at regional level, dealing with complex and interwoven issues such as agricultural developments, nature conservation, water management, spatial planning, housing and tourism, clearly illustrate the case. Policy evaluation from the perspective of policy as a control loop occasionally tends to ignore such complications.

In Chapters 3 and 4 it will gradually become clear that almost all methods of policy evaluation, and especially the most classical and robust ones, are in keeping with the first perspective: policy as a 'simple' control loop. Policy evaluation based on the perception of policy as political interaction is more complex; consequently, fewer methods, particularly standardized ones, are available.

Policy as an institutional phenomenon

The third perspective on policy further inspires the policy evaluation process, but also makes it more complex. The perceptions of policy as a control loop and as a form of political interaction primarily explore specific, often more or less separate, policy processes in the here and now. Such evaluations generally relate to the contents, the organization and implementation of a single policy intervention. There are two reasons for the emphasis in policy evaluation on the here and now, as well as the focus on processes that are regarded more or less irrespective of their context: one is political and the other is scientific in nature. First, policy evaluation research is often commissioned research and clients are mainly interested in short-term assignments dealing with specific policy efforts. Second, from a methodological point of view, it seems only natural that scientists should have a preference for policy processes that can be evaluated more or less separately since such processes are often the easiest to assess.

This approach, however, also has its limitations: real life day-to-day policy processes do not unfold in isolation, but rather as a part of a broader, relatively stable framework of policy-making, often referred to in literature as policy subsystems, policy regime, etc. – concepts that, quite rightly, express the stability or continuity of the patterns of behaviour, the styles and practices in a specific policy domain. Indeed, policy fields and domains over the years have established ways and styles of policy-making that may have become rigidly institutionalized. This is what has earlier been referred to as the pre-existing institutional context

in which a specific policy process unfolds, and by which that process, to a lesser or greater extent, is pre-structured and characterized.

Be that as it may, the institutionalized manner of policy-making is certainly related to the dominant view of one's own policy field and the central tasks

BOX 1.5 INSTITUTIONALIZATION AND POLICY

The notion of institutionalization is a basic concept in social sciences that essentially refers to the phenomenon that human behaviour gradually becomes fixed in its responses and actions. Ideals, norms and opinions, on the one hand, and practices and ways of going about things, on the other, slowly converge in broadly shared views that need no further explanation; in fixed operational methods that are accepted and adopted by all; and in rules and structures that would appear to be undisputed. Much of so-called individual behaviour is thus entrenched in 'institutional behaviour': at school, at work and at home, we display and expect others to display certain types of behavioural patterns. Although we seldom discuss them explicitly, we are well aware of the norms and rules that come with these patterns, and this gives us confidence and a sense of security. Institutions thus legitimize and stabilize social interaction.

All of the above applies equally well to policy-making. The contents and organization of policy, too, are gradually fixed in specific patterns; in common perceptions of the policy problem at hand; in conceptions of the main mission and characteristics of the policy field involved, as well as its boundaries with adjacent fields; and in accepted views of who the principal players are, what the balance of power between these players is, and how they interact. Thus, many policy fields are characterized by more or less fixed definitions of problems, by more or less fixed relationships between the actors involved, and by more or less fixed ways of policy-making and policy implementation.

Policy scientists differ in the degree to which they focus on the institutionalization of policy, how significant they find it and how they interpret both the construction and the impact of these institutional features. As a result, they refer to them in different terms. Some policy scientists emphasize the lasting predominance of historically established institutions, and focus on path dependency and other phenomena of historical institutionalism. Others emphasize the fact that institutions constantly need to be renewed as their legitimacy is challenged by new ideas and new developments. The position of national environmental policies in a gradually globalized context is such a challenging issue. Furthermore, some scholars emphasize the existence of 'discursive coalitions' to indicate the far-reaching and often lasting discursive unanimity within a policy domain field; others talk about 'regimes' to highlight regulatory aspects. It is worthwhile, for instance, evaluating to what extent the European Habitats Directive would have changed discourses and coalitions at national level in different member states. Other scholars prefer to use the concept of 'networks' to emphasize the strong connections between the actors involved; while yet others use the concept of 'institutional arrangements' to point out the close ties between policy contents and policy organization. Finally, some refer to the concept of 'policy style' to characterize the more or less fixed way of policy conduct in a specific field.

involved; of the relationships between the principal actors; of the distribution of funds, knowledge and power; of the style in which people or groups of people deal with one another; of the manner in which policy is implemented; of the relationships with local authorities or quasi-autonomous administrative bodies; etc. When a civil servant who has been transferred to a different policy field says something like: 'Water policy is, of course, completely different from waste policy', it is often these kinds of features that he or she is referring to. Also relevant in this context is the manner in which a policy field reacts to new events; to the way in which scientific input is ensured; to the way in which the interaction with the market and civil society is organized; and so on. The latter characteristics, in fact, refer to what can be labelled as organizational flexibility, which, in turn, largely explains the learning capacities of institutions and organizations. And 'learning', as we will see, could or should be a major outcome of policy evaluation.

It is hard to identify the institutional features of a policy field or subfield (e.g. the institutional characteristics of waste management, water policies or nature conservation and biodiversity) without a clear point of reference. More or less theoretically constructed typologies are helpful; but quite often scholars do carry out empirical comparisons. In general, there are essentially three ways of comparatively describing the characteristic institutionalization of a policy field: an international comparison between two similar policy fields, thus taking an international comparative perspective; a national comparison between two policy fields, often called a cross-sector approach; and a comparison between a specific policy field 20 years ago and today: a longitudinal comparison. Each of these methods is labour intensive and time consuming.

Two real life examples might clarify the case policy evaluation from an institutional perspective. The first relates to the policy evaluation question as to why organic farming, in some countries at least, has thus far achieved only a very modest market share, well below the stated policy goals. Such an evaluative question cannot be answered adequately if, from a rational actor point of view or from a policy as interaction perspective, one considers just a single or a few policy interventions: some regulatory measures, some economic incentives and a series of communicative efforts to endorse organic farming and the distribution and consumption of its products. From an institutional perspective, one requires, instead, an evaluation of the:

- institutional features of organic and traditional farming;
- institutionalized practices within the production, distribution and consumption chains;
- logistical, financial, scientific and technological support offered to the respective sectors.

In other words, such an evaluation requires an analysis of the institutional context, rather than an analysis of the impact of this or the other policy measure.

The second real life example concerns an evaluation as to why scientific research, including deliberately policy-oriented research, in the environmental field – or an associated subfield – has had such a limited impact upon environmental policy-making. This question again cannot be adequately answered on the basis of an evaluation of individual research projects and their valorization (or lack thereof). What is required is an analysis of the institutional organization of knowledge production, and how it is stimulated and applied, including the operational approaches and styles in the field, etc. What can be explored is whether traditional ways of knowledge production and use, or old ‘knowledge arrangements’, obstruct more contemporary approaches. Only this type of evaluation can yield recommendations for a reorganization of the entire policy field of environmental knowledge. And only this kind of reorganization of the context can lead to a policy conduct that brings the attainment goals closer.

Based on this kind of research, policy scientists have discovered just how stable and lasting some of these institutional patterns are, how they are continually being reproduced and consolidated, and how resistant they are to change. In certain cases, stability can turn into inertia or rigidity, and this, in turn, may leave a policy field ill equipped for tackling a new policy problem. After all, the latter is likely to necessitate an approach that does not fit in (well) with existing policy tradition or policy context. In terms of policy evaluation, if policy processes are not attuned to the context or vice versa, then the chances of achieving effective policy are greatly compromised. On the other hand, though, the environmental policy field (since it is still young compared to others, and quite constantly challenged with new issues and developments) displays rather high institutional dynamics. Air pollution during the 1970s turned into acidification in the 1980s, and into climate change and concern over particulates over the last decades. Not only do we witness major shifts in the problem formulation, these are paralleled with new policy strategies and new organizational forms, underpinned with scientific expertise from other disciplines, etc. Similar challenges are to be witnessed in fields such as nature conservation, with conservation work evolving into biodiversity policies, water policies evolving into integrated water management, corporate environmental management broadening into corporate social responsibility, etc. In all of these cases, new problem formulations create new goals, strategies and organizations, while the increasing globalization and Europeanization of these environmental domains simultaneously encourage reorganization – in other words, we witness institutional challenges here that ask for institutional responses.

It is clear to see that at the heart of this institutional perspective lies the question of what the typical features of a certain institutional context are and how they affect specific policy processes and policy products. The quintessence of an institutional evaluation is to know whether that institutional context is suitable and adequately equipped for the type of policy one intends to pursue. And the ensuing recommendations will relate to the type of institutional context that is

best or better suited for that type of policy. In other words, recommendations assume the form of an institutional design.

These questions are not easy to answer and the type of recommendations referred to are challenging to formulate. As a matter of fact, policy evaluations at an institutional level are forced to build on an array of evaluations at lower, more concrete, levels. At the same time, it is clear that they reveal structural causes of policy success or policy failure and seek out practical reasons for improving policy. But, again, the methods of institutional policy evaluation are less robust and have not been as extensively tested, so they will not be studied in great detail in the subsequent chapters.

1.2 CRITERIA FOR POLICY EVALUATION

At the beginning of this chapter, we defined policy evaluation as *a scientific analysis of a certain policy area, the policies of which are assessed for specific criteria and on the basis of which recommendations are formulated*.

In order to demonstrate that policy evaluation can be approached from several angles, the first section dealt with the question: what is policy? The answer to this question was that, starting from three different perspectives on policy and policy-making, policy evaluation may be evaluated from three different perspectives as well: as a control loop, as a form of political interaction and as an institutional phenomenon. We can easily imagine these perspectives as three levels of analysis. Which level of analysis is used obviously depends upon the policy evaluation question that is being addressed. In turn, and less obvious, is the conclusion that any policy evaluation question thus (even unintended) fits into a certain perspective – and therefore has to be assessed according to its suitability. In addition, we noted that as we move from the first perspective to the third, increasingly fewer tested methods of research are available. This is why, in Chapters 3 and 4, we focus on policy evaluation methods that approach policy as a rational goal-oriented process and as a form of political interaction. However, with regard to certain issues, perspectives for institutional evaluation will be outlined.

In this section we deal with the question: what are the criteria against which policy is to be evaluated? A short historical outline leads us to three main sets of criteria, which we subsequently relate to the internationally formulated criteria of ‘good governance’.

1.2.1. A brief history of policy evaluation and criteria

Policy evaluation, only recently gaining popularity in the environmental field, is much older than many people think. As the governance of national states grew increasingly complex during the course of the 19th century, national parliaments

found themselves barely able to carry out their task of checking the executive branch of government. They lacked the expertise to comprehend the intricate mix of government budgets, income and expenditure, and annual accounts, let alone to evaluate them. Against this backdrop, audit offices or comparable bodies were established in almost all parliamentary democracies in order to provide support to national parliaments. In some cases, they were re-established because audit offices had, in fact, been around in some European early-modern cities and kingdoms in various forms. The successive competencies and responsibilities that these audit offices held offer some excellent general insights into the development of national states, and their views on, and practices of, governance, democracy and control, leading to the development of policy evaluation.

Likewise, the successive criteria that were emphasized within the monitoring (i.e. the evaluative role of these audit offices) are extremely instructive. Traditionally, audit offices, in their capacity of special advisers to the legislative branch of government, were (and still are) responsible for monitoring whether government policy is lawful, whether the budgets are being used for the purposes they were intended for, etc. These criteria bear witness to the traditional, mainly statutory, debate on the role of the state, the state's right to levy taxes and draw up public budget, and citizens' right to protection against the arbitrariness of that same state. In this context, it is understandable that this form of 'policy evaluation' focuses strongly on mainly legally inspired criteria, such as the legitimacy of government expenditure, and that it uses these criteria as crucial elements for evaluating quality of policy. Hence, many scholars refer to this type of policy evaluation as 'judicial evaluation'.

During the post-war era, especially as a result of large-scale public investment programmes in the US, including on defence and space exploration, the need arose to scrutinize and evaluate government policy in a more economics-based way. The notion that perhaps government policy should be managed and evaluated in the same way as company policy gained ground, giving rise, during the 1960s and 1970s, to the concept of 'public management'. It was in the spirit of this era that the notion of policy evaluation as we know it today originated. Policy was expected to focus on solving problems, a train of thought which we have identified in section '1.2 Criteria for policy evaluation' as the rational-economic perspective on policy-making, an approach that prevailed during those years. From this point of view, policy evaluation would determine whether and to what extent the goals of a certain policy had been attained, whether those goals were actually attained through the policy intervention so that it could be said to be effective, and whether that policy was being pursued in the most appropriate or most efficient way. This approach led to the development and gradual sophistication of famous methods such as the Programme Planning and Budgeting System (PPBS), which were used for policy design and policy evaluation simultaneously. Apart from the policy domains characterized by huge public investments, such as defence and space exploration in the US, other policy

fields that have been subjected to evaluative studies internationally are most often those of employment and education. In both cases, public authorities have a particular interest in the effectiveness of specific programmes targeting vulnerable or underprivileged categories. As will become apparent in Chapters 3 and 4, many methods of policy evaluation were created and developed specifically in those policy fields. Before that, Chapter 2 will consider some of the methodological problems that are typically associated with the evaluation of environmental policies.

With the advent of neo-liberalism and new public management during the 1980s and 1990s, (business) economics became even more important as the angle from which to evaluate public policy. Performance increasingly came to be seen as the key criterion. In fact, this was not just the case in policy evaluation, but also in policy development: the performance of government – or of the kind of autonomous administrative bodies that were increasingly entrusted with policy implementation – is gauged on the basis of indicators; the output to be generated is laid out in contracts; and subsidizing is made dependent upon the actual delivery of that output. Thus, again, part of the newly emerged business management monitoring systems and evaluation processes had been incorporated within policy development. This is why scholars refer to this type of evaluation either as ‘administrative’, alluding to its earlier stages, or as ‘economic’ or even ‘managerial’, with reference to its later and current stages.

During the course of the 1980s, however, other evolutions would unfold. In addition to effectiveness, the notion of legitimacy of government intervention increasingly came to the fore. Consequently, policy evaluation gradually came to apply such political evaluation criteria as legitimacy, acceptance, participation and responsiveness. In fact, analysing and evaluating the public support or legitimacy of policy almost developed into an independent field of research. Moreover, policy evaluation, because of its focus on public support and legitimacy, was increasingly shifting towards more participatory approaches. These approaches provided not only for the evaluation of the actual actors involved, but also for evaluation based on their objectives, wants and preferences with regard to content, as well as to the policy process and organization. It won't come as a surprise that this type of evaluation is most often referred to as ‘political evaluation’. This type of evaluation is mostly carried out in close cooperation with policy-makers and other stakeholders in order to make policy evaluation as much a multi-actor activity as policy-making itself. Examples of participatory policy evaluation approaches are provided in Chapter 4.

As the latter set of criteria suggests, a gradual shift had taken place in the perspective adopted on policy evaluation: whereas policy evaluation was originally directed almost exclusively at administrative monitoring and economic accountability, it slowly also came to focus on aspects such as ‘learning’, reflection on the various policy mechanisms and generating responses to them that would improve policy, make it more effective, enhance its legitimacy, and so on. Over

the last decades or so, ‘learning’ has become one of the buzzwords in organizational theory and in public policy studies, and some suggest learning to be the predominant contribution that policy evaluation can provide and/or should be aimed at. While we agree, in general terms, about the desirability of (organizational, social, governmental) learning, especially when facing very complex and uncertain environmental issues, one cannot consider ‘learning’ to be a more or less automatic (side) effect of policy evaluations. Even if evaluations are deliberately targeting learning, we know from literature on knowledge utilization, organizational change and institutional change that organizations and government bodies do differ substantially in their willingness and their capacity for learning. These characteristics, in turn, are related to features such as rigidity versus flexibility, openness for new information, and hierarchy versus horizontal styles of working, which are reported on in depth in policy studies. Policy evaluation, therefore, even though it comes up with high-quality research and with highly relevant recommendations, does not guarantee any learning. Nevertheless, some approaches in policy evaluation, particularly those emphasizing the importance of participation, also pay a great deal of attention to (increasing) the learning effect, as Chapter 4 will clarify in more detail.

1.2.2 The JEP triangle: The three rationalities of policy evaluated

Various policy scientists have pointed out that public policy is based on more than a single rationale or basic motive. Generally speaking, three underlying rationales of public policy are distinguished: a juridical or judicial rationale, an economic rationale (sometimes referred to as a business rationale) and a political rationale. Together these are known as the JEP triangle. Without presuming a one-to-one relation, these three rationales more or less reflect the three types of evaluation that we sketched above as they emerged throughout the history of policy evaluation. We will briefly elucidate each of these rationales because they also form the bases of three sets of criteria used for (environmental) policy evaluation.

The juridical or judicial rationale is related to the rule of law and to the principles of good governance. As stated above, the traditional criteria of policy evaluation by an audit office are in line with these criteria. The main issues of concern here are the protection of basic rights, the principles of lawfulness and the maintenance of law and order etc.: a set of basis principles and rulings that grew in number and international attention as time went on. These lay down a number of principles regulating the rules of play among public authorities (e.g. the principles of subsidiarity, policy hierarchy, multilevel governance, etc.) and between the public authorities and civil society (e.g. the principles of legal protection, etc).

In contrast, the economic or business rationale focuses mainly on policy goal attainment, effectiveness and efficiency. Economy, performance, enforceability,

feasibility, etc. are all important focal points in this respect. Many of today's management tools for governance and its monitoring are in line with these criteria.

The political rationale behind policy refers to the principles at the core of democracy: representation; mandate and accountability; accessibility, openness and responsiveness; transparency and participation; etc. This rationale finds expression in various political rights and liberties, in the separation of powers (*trias politica*), in limitations on government power and so on, but also in the quest for legitimacy by policy-makers.

These three underlying rationales of policy are (even though their respective significance and importance may vary over time) invariably present in an ongoing political debate. At the same time, they help to distinguish three types of policy evaluation, as sketched above. And, more importantly for this section, these three rationales provide us with three sets of policy evaluation criteria. These criteria, clustered into typically juridical or judicial, managerial and political criteria, respectively, have given rise to the so-called JEP triangle (see Figure 1.2).

First and foremost, this figure demonstrates once again which criteria may be applied from each of the three perspectives. Of course, the criteria mentioned above and those described in Figure 1.2 need to be further 'operationalized' (i.e. they need to be expressed in terms of variables). Second, by juxtaposing the three sets of criteria in a triangle, it becomes clear that not only are the three rationales behind policy or the three clusters of policy evaluation criteria different, but they are also contradictory. The well-known dilemma of policy-makers, who must strike a balance between power and legitimacy, can now be translated in terms of the opposing criteria of economic efficiency, on the one hand, and political

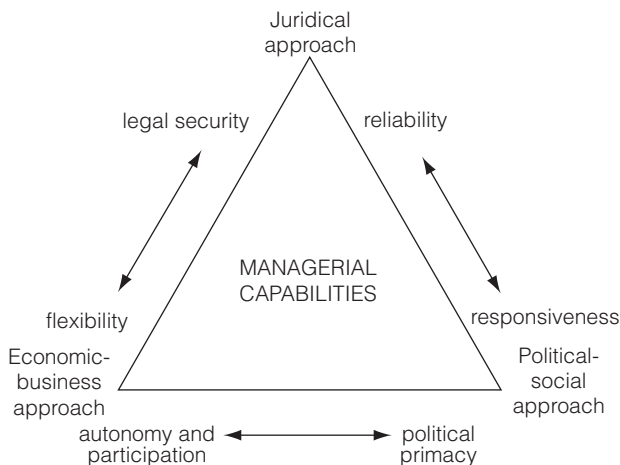


Figure 1.2 *The so-called juridical, economic and political criteria (JEP) triangle*

legitimacy, on the other. There is also an element of contradiction between the judicial criterion of legal security and the managerial criterion of flexibility. By highlighting these contradictions, the triangle also makes clear that any evaluation that takes into account only one perspective tends to be one sided. In other words, a balanced policy evaluation must encompass criteria from more than one perspective.

This, once again, suggests that an evaluation of effectiveness and efficiency should, in any case, be complemented with an assessment from other angles. Moreover, and coming back to the three perspectives set out in the previous section, Table 1.1 makes clear that ‘effectiveness’ may, indeed, have quite different meanings depending upon the perspective that one uses in policy evaluation.

Without further elaborating upon Figure 1.2 and Table 1.1, it is clear that an evaluation that focuses on managerial effectiveness and efficiency exclusively, however important these aspects may be, runs the risk of not shedding sufficient light on other evaluation criteria and other perspectives. It may, indeed, neglect certain qualities and inconveniences of the policy concerned and thus also ignore opportunities for learning and improving that particular policy.

1.2.3 Criteria of good governance?

For some time now, the United Nations has been promoting a number of principles of good governance – in this case, good public governance. The eight principles of good governance are rule of law; consensus-building; participation; responsiveness; transparency; accountability; equitability and inclusiveness; and effectiveness and efficiency (see Figure 1.3).

Table 1.1 *Effectiveness from three perspectives*

| <i>Perspective</i> | <i>Policy making is ...</i> | <i>Guiding metaphors</i> | <i>Success is ...</i> |
|--------------------|---|--|---|
| Rationalism | Problem-solving oriented | Policy cycle (regulatory cycle in engineering) | Problem-solving |
| Interaction | Interaction between agencies and interests involved | Networks, stakeholders, negotiation, deliberation | Network constitution, exchange and convergence |
| Institutionalism | Functioning of regimes; institutional arrangements | Path dependency versus institutional change (transition) | Institution-building; institutional (change) capacity |



Figure 1.3 *The principles of good governance*

Source: UNESCAP (2008)

From an historical point of view, these eight principles may be understood to be the sum total of traditional judicial principles of appropriate governance, complemented by more recent managerial or economic, as well as political, principles. At the present juncture, these eight criteria constitute a kind of benchmark for good public governance. Their purpose, especially in a global context, is both to mobilize and to programme: they present states, their leaders and their citizens with an ideal in the hope that they will aspire to it. Additionally, these principles or criteria, which were already being used by critical non-governmental organizations (NGOs), are now also increasingly relied on by organizations such as the World Bank and the International Monetary Fund to evaluate the general quality of government and, thus, the credit worthiness of states. In this way, good governance has become a basic requirement for a country to be admitted to the international community. Of course, in this area, too, pragmatism sometimes trumps principle. The same holds for the notion of corporate governance: the management of mainly internationally operating companies is also increasingly evaluated on the basis of a multitude of criteria. Traditional evaluation criteria of business economics (important primarily to shareholders) have been complemented with aspects such as social entrepreneurship and sustainability (important primarily to stakeholders).

However, we are not concerned here with the actual application and impact of such principles. We simply want to demonstrate that the eight principles of good governance, frequently referred to in domains related to globalization, development, poverty, sustainability and environment, are, in fact, a temporary

synthesis of decades, if not centuries, of policy evaluation and criteria development. Traditional legal criteria of lawfulness have been supplemented with recent political criteria such as participation, and with strongly managerially inspired criteria such as accountability. Contemporary policy evaluation will, in one way or another, need to strike a balance when applying these criteria. In Chapter 3 and 4 it will become apparent how various methods of policy evaluation take this into account.

REFERENCES

As already mentioned, a full reference list as is the case in most classical textbooks, is not provided here. For detailed references to the various methodological approaches, we refer to Chapters 3 and 4. What follows is a handful of books on policy evaluation that inspired us when preparing courses, when carrying out research in policy evaluation and when conceiving this book.

- Arts, B. and Goverde, H. (2006) 'The governance capacity of (new) policy arrangements: A reflexive approach', in B. Arts and P. Leroy (eds) *Institutional Dynamics in Environmental Governance*, Springer, Dordrecht, The Netherlands
- Bovens, M., t'Hart, P. and Peters, B.G. (eds) (2001) *Success and Failure in Public Governance: A Comparative Analysis*, Edward Elgar, Cheltenham, UK
- Clarke, A. and Dawson, R. (1999) *Evaluation Research: An Introduction to Principles, Methods and Practice*, Sage, London
- Dunn, W. N. (2003) *Public Policy Analysis: An Introduction*, 3rd edition, Prentice-Hall, Englewood Cliffs, NY
- Edelenbos, J. and van Eeten, M. (2001) 'The missing link: Processing variation in dialogical evaluation', *Evaluation*, vol 7, no 2, pp204–210
- EEA (European Environment Agency) (2001) *Reporting on Environmental Measures: Are We Being Effective?*, EEA, Copenhagen, Denmark
- Fischer, F. (1995) *Evaluating Public Policy*, Nelson-Hall, Chicago, IL
- Furubo, J. E., Rist, R. C. and Sandahl, R. (eds) (2002) *International Atlas of Evaluation*, Transaction Publishers, New Brunswick, NJ
- Hage, M. and Leroy, P. (2007) *Stakeholder Participation Guidance for the Netherlands Environmental Assessment Agency: Main Document/Practice Guide/Check List*, MNP/RU Nijmegen, www.mnp.nl/en/publications/2008/StakeholderParticipation_Guidance_for_theNetherlandsEnvironmentalAssessmentAgency_MainDocument.html, accessed 1 March 2008
- Hall, P. (1993) 'Policy paradigms, social learning and the state: The case of economic policy making in Britain', *Comparative Politics*, vol 25, no 3, pp275–296
- Howlett, M. and Ramesh, M. (2003) *Studying Public Policy – Policy Cycles and Policy Subsystems*, Oxford University Press, Oxford, UK

- Leeuw, F., Rist, R. C. and Sonnischen, R. C. (eds) (1994) *Can Governments Learn? Comparative Perspectives on Evaluation and Organisational Learning*, Transaction, New Brunswick, NJ
- Madaus, G. F., Scriven, M. and Stufflebeam, D. L. (eds) (1986) *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Kluwer-Nijhoff, Boston, MA
- Owen, J. M. and Rogers, P. (1999) *Program Evaluation: Forms and Approaches*, Sage, London
- Patton, M. Q. (1997) *Utilization-Focused Evaluation: The New Century Text*, Sage, Thousand Oaks, CA
- Pressman, J. L. and Wildawsky, A. B. (1984) *Implementation: How Great Expectations in Washington Are Dashed in Oakland*, University of California Press, Berkeley, CA
- Sabatier, P. A. and Jenkins-Smith, H. C. (eds) (1993) *Policy Change and Learning: An Advocacy Coalition Approach*, Westview Press, Boulder, CO
- Stufflebeam, D. L. (2001) 'Evaluation models', *New Directions for Evaluation*, no 89
- UNESCAP (United Nations Economic and Social Commission for Asia and the Pacific) (2008) 'What is good governance?', www.unescap.org/pdd/prs/projectactivities/ongoing/gg/governance.asp, accessed 1 March 2008
- Vedung, E. (2000) *Public Policy and Program Evaluation*, Transaction Publishers, New Brunswick, NJ
- Worthen, B. R., Sanders, J. R. and Fitzpatrick, J. L. (1997) *Program Evaluation: Alternative Approaches and Practical Guidelines*, Longman Publishers, New York, NY