As politicians know only too well but social scientists too often forget, public policy is made of language. Whether in written or oral form, argument is central in all stages of the policy process. Discussion goes on in any organization, private or public, and in any political system, even a dictatorship; but it is so much at the heart of democratic politics and policy that democracy has been called a system of government by discussion. Political parties, the electorate, the legislature, the executive, the courts, the media, interest groups, and independent experts all engage in a continuous process of debate and reciprocal persuasion.

This process, as liberal theorists from John Stuart Mill and Walter Bagehot to Lord Lindsay and Ernest Barker have described it, begins with expressions of general concerns and ends in concrete decisions. Each stage of deliberation has its own function and its own organ. Parties identify issues and formulate programs; the electorate discusses issues and candidates and expresses a majority in favor of one of the programs; the legislative majority translates programs into laws, in constant debate with the opposition; finally, the discussion is carried forward to the chief executive and the cabinet, where it is translated into specific policies. Each of the stages and organs of public delib-
eration is independent, but only within the limits, and as a part, of the entire process: "the free and sovereign thing is the whole process of discussion."

This is an idealized model of democratic policy-making. It overlooks the play of power and influence, the uneven distribution of knowledge, the low level of active citizen participation, and many other factors that figure prominently in modern theories of public policy. But it emphasizes something that these theories have neglected—the extraordinary potential of persuasion and the centrality of two-way discussion to democracy.

Every politician understands that arguments are needed not only to clarify his position with respect to an issue, but to bring other people around to this position. Even when a policy is best explained by the actions of groups seeking selfish goals, those who seek to justify the policy must appeal to the public interest and the intellectual merits of the case. Perhaps these are only rationalizations, but even rationalizations are important since they become integral parts of political discourse. We miss a great deal if we try to understand policy-making solely in terms of power, influence, and bargaining, to the exclusion of debate and argument.

Argumentation is the key process through which citizens and policymakers arrive at moral judgments and policy choices. Public discussion mobilizes the knowledge, experience, and interest of many people, while focusing their attention on a limited range of issues. Each participant is encouraged to adjust his view of reality, and even to change his values, as a result of the process of reciprocal persuasion. In this way, discussion can produce results that are beyond the capabilities of authoritarian or technocratic methods of policy-making.

constraints make such an experiment practically impossible. Similarly, the choice of a dose-response function to determine the “virtually safe” dose of a toxic substance must be treated at present as a trans-scientific question. There are literally thousands of mathematical functions that fit the experimental data equally well, but no firm scientific basis now exists for choosing among the different possibilities. However, the choice of a particular function has a major effect on regulatory decisions.

When science, technology, and public policy intersect, different attitudes, perspectives, and rules of argument come into sharp conflict. Scientific criteria of truth clash with legal standards of evidence and with political notions of what constitutes sufficient ground for action. Factual conclusions are not easily separable from considerations having to do with the plausibility of the opponent’s assumptions and his selection of the evidence or choice of methodology. And because there seems to be no objective way of checking the conclusions of analysis, the credibility of the expert becomes as important as his competence.

Increasingly, public debates about regulatory decisions, nuclear safety, technology assessment, and similar trans-scientific issues tend to resemble adversary proceedings in a court of law, but with an important difference—the lack of generally accepted rules of procedure. Some participants are able to take advantage of the relative informality of the process, but to scientists even codified adversary procedures seem inappropriate and alien to their tradition. In science the issue is not a witness’s credibility but his specific competence—his ability to establish scientific truth—and this is not reliably established by an adversary debate. Hence various proposals to resolve disputes about scientific issues with policy implications by carefully breaking down a problem into purely technical and purely political components. Experts should deal only with the technical issues, turning their evaluations over to the political process for determination of the appropriate policy response.

For example, the “science court” proposed by Arthur Kantrowitz would only examine, and decide upon, questions of scientific fact. After the evidence has been presented, questioned, and defended, the panel of judges (who are established experts in areas adjacent to the dispute) issue a report in which the points of agreement among the experts are given. The report may also suggest specific research projects to clarify points that remain unsettled.

But how can one separate the scientific from the political and value components of policy issues that encompass both? And if trans-scientific questions do not come within the purview of the scientific court, why use a quasi-judicial procedure? If the question is unambiguously scientific, then the procedures of science rather than quasi-legal ones are appropriate. Where the issues cannot be settled with existing scientific knowledge or from research that could be carried out reasonably rapidly and without excessive expense, then the answers must be trans-scientific and an adversary procedure that involves both experts and generalists seems the best alternative.

Dialectical confrontation between generalists and experts often succeeds in bringing out unstated assumptions, conflicting interpretations of the facts, and the risks posed by new projects. Technical experts are naturally biased in the assessment of their proposals and are more likely to be skeptical of any evidence of possible adverse effects than someone less committed to that particular project. The initial assumption is that the innovation will achieve what the innovator claims for it and that it will have no negative consequences that could reduce the attractiveness of its practical implementation. For example, the consciousness of the dangers inherent in nuclear engineering in the United States and Western Europe is largely the result of public debate. Where nuclear technology has been allowed to develop according to its own logic, unhampered by criticism and public concern, as in the Soviet Union, it has produced few of the safety features (such as containment shells for pressurized water reactors) that are now standard in the West.

Thus, technological expertise cannot be relied upon to discover the characteristic risks and the social implications of new

technologies. The essential need today is an improvement in the methods and conditions of critical debate and their institutionalization at all levels of policy-making. Actually, attempts to develop methods of critical inquiry adapted to the process of public deliberation go back to the origins of democracy.

Building on the practice of government by discussion in the city-state, the Greeks developed a general technique of critical discourse which they called dialectic. This is a method of argumentation characterized not so much by the form of reasoning (though discussion by question and answer came to be regarded as its paradigmatic form) as by the nature of its premises and the social context of its applications. Logic and mathematics start from axioms, or propositions deduced from axioms, while the premises of dialectic are merely plausible. The starting point of a dialectic argument is not abstract assumptions but points of view already present in the community; its conclusion is not a formal proof, but a shared understanding of the issue under discussion; and while scientific disciplines are specialized forms of knowledge, available only to the experts, dialectic can be used by everyone since, as Aristotle put it, we all have occasion to criticize or defend an argument.

For the Greeks dialectic had three main uses. First, as a method of critical inquiry into the foundations and assumptions of the various specialized disciplines. Second, as a technique for arguing in favor of one's own opinions and a procedure for clarifying controversial issues. Finally, as an educational process that transforms the common man into an informed citizen and the specialist into a person able to communicate with his fellow citizens.

This ancient notion of dialectic is quite relevant to our inquiry into the role of analysis in public deliberation. In fact, it seems to capture the essential elements of that role much better than the stereotyped characterization of policy analysis found in current textbooks. Like dialectic, policy analysis usually starts with plausible premises, with contestable and shifting viewpoints, not with indisputable principles or hard facts. Like dialectic, it does not produce formal proofs but only persuasive arguments. The key problem facing both dialecticians and analysts is how to base plausible inferences on values or opinions when hard facts are not available. Finally, policy analysis, like dialectic, contributes to public deliberation through criticism, advocacy, and education. Good policy analysis is more than data analysis or a modeling exercise; it also provides standards of argument and an intellectual structure for public discourse. Even when its conclusions are not accepted, its categories and language, its criticism of traditional approaches, and its advocacy of new ideas affect—even condition—the policy debate.

THE ARGUMENTATIVE FUNCTION OF POLICY ANALYSIS

The purpose of this book is to discover the main implications of a dialectic conception of policy analysis. In it I attempt to develop a single idea: the notion that in a system of government by discussion, analysis—even professional analysis—has less to do with formal techniques of problem solving than with the process of argument.

The job of analysts consists in large part of producing evidence and arguments to be used in the course of public debate. Its crucial argumentative aspect is what distinguishes policy analysis from the academic social sciences on the one hand, and from problem-solving methodologies such as operations research on the other. The arguments analysts produce may be more or less technical, more or less sophisticated, but they must persuade if they are to be taken seriously in the forums of public deliberation. Thus, analysts, like lawyers, politicians, and others who make a functional use of language, will always be involved in all the technical problems of language, including rhetorical problems.

Rhetoric is the craft of persuasion, the study of all the ways of doing things with words. The Athenians used to make annual sacrifices to the goddess of persuasion (Peitho) in recognition of the extraordinary power of language. Today persuasion is often

regarded as a dishonest or merely “rationalizing” use of arguments; it is propaganda, brainwashing, manipulation of public opinion. Persuasion can indeed be used in these ways. But in free debate, persuasion is a two-way interchange, a method of mutual learning through discourse. Real debate not only lets the participants promote their own views and interests, but also encourages them to adjust their views of reality and even to change their values as a result of the process.

A persuasive argument is not a logical demonstration, but it does not become irrational or mere rationalization because of this. Most value judgments are formed in persuasive interchange. To reduce reason to logical calculation and proof about whatever does not matter enough to engage commitment is, as Wayne C. Booth has written, to create a torn picture of the world, with all our values on one side and all our rational faculties on the other. Since to say anything of importance in public policy requires value judgments, this artificial separation between values and rational capacities is a threat to all notions of public deliberation and defensible policy choices.

As I will show in the next chapter, even technical policy analysts cannot dispense with persuasion. On the one hand, facts and values are so intertwined in policy-making that factual arguments unaided by persuasion seldom play a significant role in public debate. On the other hand, persuasion is needed in order to increase both the acceptability of advice and the willingness to act on less than conclusive evidence. To explain and defend a reasonable course of action under circumstances where the theoretical optimum is either unknown or practically unattainable is an essential part of the analyst’s job.

Feasibility analysis, to be discussed in chapter 4, is perhaps the best illustration of the necessary interplay of empirical and persuasive arguments. Fashioning mutual understandings about the boundaries of the possible in public policy is arguably the most important contribution that analysts can make to public debate. However, calculating optimal or second-best solutions within given constraints is only the static part of feasibility analysis; the dynamic and more important part is discovering means to push out the boundaries of the possible. Doing this requires both objective analysis and persuasion: what is possible often depends on what the political system considers fair or acceptable. Many policy constraints can be eased only by changing attitudes and values; as already noted, this always involves a certain amount of persuasion.

ARGUMENTATION AND EVALUATION

Persuasive arguments play an even larger role in evaluative discourse. Whenever new evaluative criteria or a reform of old criteria are being considered, it is open to anyone to put forward a proposal as to what the criteria should be and to use persuasion in order to influence others to accept the proposal.

The characteristic difficulty of policy evaluation is precisely the multiplicity of admissible standards. Citizens, legislators, administrators, judges, experts, the media—all contribute their particular perspectives and criteria. This variety of viewpoints is not only unavoidable in a pluralistic society; it is also necessary to the vitality of a system of government by discussion. Nevertheless, as Northrop Frye has remarked in the context of literary criticism, there seems to be no reason why the larger understanding of public policy to which these separate perspectives are contributing should remain forever invisible to the different evaluators, as the coral atoll is to the polyp.

Multiple policy evaluation should also be possible. It would recognize the legitimacy of the different perspectives but would also seek—by making these perspectives more aware of one another—to reach a level of understanding and appreciation that is more than the sum of the separate evaluations. The purpose is not to construct a grand model that would combine all the partial perspectives into one general criterion of good policy—a weighted average, as it were, of equity, effectiveness, legality, and any other relevant standard—but to contribute to a shared understanding of the multiple perspectives involved.

Evaluation will be discussed in detail in chapter 8. Here I will mention only one aspect of the subject—the evaluation of analysis and other types of research with policy implications. The assessment of policy arguments, like the assessment of scientific or legal arguments, necessarily involves formalities. When the issues under discussion require complex patterns of reasoning and large amounts of data of doubtful reliability and relevance, explicit rules of evidence become particularly important. A good example is the judicial law of evidence with its sophisticated distinctions among proofs of facts, testimony, hearsay, presumptions, interpretations, and other sources of information.

In chapter 3 I introduce a number of distinctions (for example, among data, information, and evidence) whose main purpose is to facilitate the evaluation of policy arguments. The importance of drawing distinctions that are usually overlooked in conventional treatments of policy analysis can be illustrated with reference to the categories “evidence” and “argument.” The argument is the link that connects data and information with the conclusions of an analytic study. The structure of the argument will typically be a complex blend of factual statements and subjective evaluations. Along with mathematical and logical deductions it will include statistical, empirical, and analogical inferences, references to expert opinion, estimates of benefits and costs, and caveats and provisos of different kinds. This unavoidable complexity makes any direct, informal testing of the argument quite impossible. Whatever testing is done must rely on a variety of standards that depend on the analytic methods employed, on the plausibility and robustness of the conclusions, and on agreed-upon criteria of adequacy and effectiveness.

The nature of the evidence is crucial in this kind of testing, since an incorrect assessment of its strength and suitability before it is included in the argument can lead to pitfalls in drawing conclusions. Evidence is not synonymous with data or information. It is information selected from the available stock and introduced at a specific point in the argument in order to persuade a particular audience of the truth or falsity of a statement. Selecting inappropriate data or models, placing them at a wrong point in the argument, or choosing a style of presentation that is not suitable for the intended audience, can destroy the effectiveness of information used as evidence, regardless of its intrinsic cognitive value. Thus, criteria for assessing evidence are different from those used for assessing facts. Facts can be evaluated in terms of more or less objective canons, but evidence must be evaluated in accordance with a number of factors peculiar to a given situation, such as the specific nature of the case, the type of audience, the prevailing rules of evidence, or the credibility of the analyst.

Disciplines like history and law, which depend on information that cannot automatically be assumed to be reliable or relevant, explicitly recognize evidence as an autonomous conceptual category. Policy analysis, too, often involves large amounts of data of doubtful reliability and relevance, but problems of evidence have not received the same attention here.

For instance, according to a view that is widespread among analysts, a good policy model should resemble as much as possible the formalized models of the more successful “hard” sciences. Accordingly, there is a dangerous tendency to regard model outputs as facts, rather than as evidence to be used in an argument together with other data and information. As a result, “the documentation of models and source data is in an unbelievably primitive state. . . Poor documentation makes it next to impossible for anyone but the modeler to reproduce the modeling results and probe the effects of changes to the model. Sometimes a model is kept proprietary by its builder for commercial reasons. The customer is allowed to see only the results, not the assumptions.”

Such gross disregard for the most elementary rules of evidence is a direct consequence of the failure to recognize the crucial argumentative aspect of policy analysis. In turn, this failure can be explained by the adherence of most analysts to a methodology that is more concerned with what decisions are made than with how they are made, or how they may be justified in the forums of public deliberation.

DECISIONISM

The image that lies behind this methodology has been called decisionism—the "vision of a limited number of political actors engaged in making calculated choices among clearly conceived alternatives."10 An actor's choices are considered rational if they can be explained as the choosing of the best means to achieve given objectives. In this view the economic model of choice becomes the appropriate paradigm for all policy problems.

For example, a well-known textbook on policy analysis introduces its subject matter as follows: "How choices should be made—the whole problem of allocating scarce resources among competing ends—is the stuff of economics and the subject of this book."11 Similar statements can be found in the writings of influential authors like Hitch, McKean, Enthoven, and Quade.12

In order to decide rationally the policymaker must specify his objectives; lay out the alternatives by which the objectives may be accomplished; evaluate the consequences of each alternative; and choose the action that maximizes net benefits. If the recipe sounds familiar it is because the logical structure of allocative decisions is the same whether the decisions are taken by individual consumers, by private entrepreneurs, or by public managers and policymakers. Hence the appeal to a generalized logic of choice which decisionists found ready-made in microeconomics and decision theory. Moreover, since the logic of choice has been investigated primarily in the context of market transactions, some writers have argued that the main, if not the only, object of policy analysis is to extend the principles of rational choice from the sphere of private economic transactions to that of public policy-making. In fact, rational policy-making, decision making, problem solving, and policy analysis become nearly synonymous terms. For example, the recipe for maximizing net benefits may be interpreted either as a description of ideal policymaking or as a prescription for policy analysis. The underlying notion of rationality is the same in both cases: rationality is maximizing something, choosing the best means to a given end.

The view of policy analysis as decision theory "writ large" has considerable intuitive appeal and provides a useful way of formulating a variety of practical problems: whether to use a particular vaccine to halt the spread of a threatened epidemic; where to build a dam; how to reduce the response time of the fire department of a big city. Not surprisingly, these or similar examples are the standard illustrations used in conventional textbooks.

The decisionist approach was developed during the Second World War and was given emphasis and formal statement in the early 1950s at the Rand Corporation and other policy-oriented think tanks. It is a conceptual compound that includes elements from operations research and management science, from microeconomics and decision theory, and a dash of social and behavioral science. A continuous line of development runs from the wartime studies of military operations, of logistics and tactics, to the early industrial applications of new quantitative methods, to systems analysis, and then to policy analysis. Technical efficiency as a goal or criterion of choice has been replaced by economic efficiency, which in turn has been tempered by considerations of equity and political feasibility. But the original analytic framework is still clearly recognizable.

The early practitioners of this approach claimed to be able to give useful advice by applying scientific methods of analysis to data collected from actual operations. In fact the situations investigated by operations researchers during the war fit the natural science paradigm rather well. Military operations could be regarded as representative of a class of repetitive situations where models built up in response to earlier examples of the

situation could be checked against later examples, monitored while proposals for improved actions were in use, and used to detect their own dwindling validity as the situation changed.13

An important characteristic of early studies of military and industrial operations was a reasonable clarity in the definition of the role of analysts and decision makers. Whether the users of analysis were high-level military officers or high-level managers, analysis was done primarily or even (because of the requirement of military or industrial secrecy) exclusively for them. The analyst did not have to address any audience other than the decision maker, or a small group of decision makers, who had commissioned the study. Problems of communication and implementation could be safely assumed to be the responsibility of a well-defined hierarchical authority, and the same authority would ensure legitimacy and provide criteria of quality and effectiveness.

By the 1960s, however, the nature of the problems analysts were investigating, and the organizational and political context in which they operated, had radically changed. The problems claiming analytic attention were becoming broader and more complex. Strategic, rather than tactical, issues loomed increasingly important, while subjective uncertainty was seen to be more crucial than the statistical regularities assumed in earlier models. At the same time, the growing role of analysis in public debates meant that analysts—no longer discreet advisors to the prince but actors in a political process in which advocacy and persuasion could not be neatly separated from objective analysis—had to pay attention to questions of equity and political feasibility.

In the early 1970s policy analysis came to replace systems analysis as the professional label denoting the activity of analysts who were concerned with public issues. This terminological change was meant to suggest a synthesis of the conflicting logics of economic and political rationality. In practice, since political science seemed unable to provide a set of concepts and analytic techniques comparable to the strong normative structure of microeconomics, the majority of policy analysts remained firmly committed to a decisionist methodology.

THE LIMITATIONS OF DECISIONISM

The limitations and biases of the decisionist approach are perhaps less obvious than its merits, but they affect almost every aspect of the teaching and practice of policy analysis. However, I shall restrict my critical remarks to a few points of particular relevance for my subsequent discussion. The purpose is to illustrate the kinds of issues and arguments that this approach tends to exclude from analysis.

To begin with, the decisionist approach assumes a unitary decision maker, or a group acting as a unit, and is not immediately applicable to situations involving two or more actors with different objectives. The model of rational choice that underlies this approach has been developed for an individual who wishes to be consistent and expresses this consistency in the way he orders his preferences and evaluates the probabilities of uncertain events. When several individuals are involved, the model does not require them to agree on their orderings and evaluations; each may be rational (that is, consistent) in holding quite divergent views. If a joint decision is required, they will have to resolve their differences through interactive processes like negotiation and persuasion, about which the model is silent.

A fortiori, this methodology ignores conflicts between the interests and perceptions of different government agencies. But whenever such conflicts are present, important questions arise about the appropriate assumptions regarding the behavior of other public agencies in the formulation of policy by any particular agency. As I argue in chapter 6, all policy instruments are effectively constrained within certain ranges by political and administrative considerations. Therefore, it is important for policymakers to know which variables are in fact within their control and to what extent, and in this respect a unitary model of policymaking is not very useful.

Another key assumption is that there is no essential distinction

between policies and decisions, so that all policy problems can be discussed in the language of decision making. But as Philip Selznick has pointed out,

decisionmaking is one of those fashionable phrases that may well obscure more than it illuminates. It has an air of significance, of reference to important events; and the mere use of the phrase seems to suggest that something definite has been scientifically isolated. But decisions are with us always, at every level of experience, in every organism. The general features of all choice, or of all social choice, may some day be convincingly stated. But it will still be necessary to distinguish the more and the less trivial; and, if there is any order in this phenomenon, to identify some kinds of decisions, linking them to the distinctive problems or situations out of which they arise.14

In fact, contrary to a widespread belief, decision theory does not apply to decision making in general, but only to choice situations of a rather special type. The decision of decision theory is a choice that must be made in the situation immediately confronting the decision maker, taking into consideration the probable consequences of each possible course of action in the present situation. Future benefits, for example, must be defined in terms of the way they are assessed today, even though there is no reason to assume that this will coincide with the assessment of those benefits in the future.15

Confronted with two alternatives, the decision maker of decision theory chooses as if the benefits from one were greater than (or at least equal to) the benefits from the other in the present situation. As John Tukey points out, this choice does not assert anything about the actual state of affairs, or about the consequences in other situations of acting as if the chosen alternative were the best of the available ones. Decisions to act in this way are attempts to do as well as possible in specific situations, to choose wisely among the available gambles.16 Decision theorists are fond of quoting Blaise Pascal's "il faut parier, il faut choisir" (one has to bet, one has to choose), and problems connected with games of chance have supplied the original paradigm for probability theory and its modern offspring, decision theory.

There are situations, in private life as well as in business and in government, where individuals do in fact choose under the conditions covered by the theory. But in many other situations the theory is not particularly useful, either prescriptively or descriptively.

Thus, important policy decisions are more than attempts to do as well as possible in the situation immediately confronting the policymaker. They are taken after careful deliberation and are judged by their long-run effects rather than by their immediate consequences. Once taken, they are retained for some time, providing direction and consistency to the various activities of government. As a former presidential adviser puts it, "Most presidential decisions are too far-reaching and too irrevocable to be taken in haste, when the facts are uncertain, when the choices are unclear, or when the long-range consequences are not as discernible as the immediate reactions and results."17 What Theodore Sorensen describes here are not decisions in the sense of decision theory, but policy judgments that are more usefully discussed in the language of prudential discourse.

A third limitation of decisionism is its exclusive preoccupation with outcomes and lack of concern for the processes whereby the outcomes are produced. A lack of concern for process is justified in some situations. If the correctness or fairness of the outcome can be determined unambiguously, the manner in which the decision is made is often immaterial; only results count. But when the factual or value premises are moot, when there are no generally accepted criteria of rightness, the procedure of

decision-making acquires special significance and cannot be treated as purely instrumental.

Even in formal decision analysis the explicit recognition of uncertainty forces a significant departure from a strict orientation toward outcomes. Under conditions of uncertainty different alternatives correspond to different probability distributions of the consequences, so that it is no longer possible to determine unambiguously what the optimal decision is. Hence, the usual criterion of rationality—according to which an action is rational if it can be explained as the choosing of the best means to achieve given objectives—is replaced by the weaker notion of consistency. The rational decision maker is no longer an optimizer, strictly speaking. All that is required now, and all that the principle of maximizing expected utility guarantees, is that the choice be consistent with the decision maker’s valuations of the probability and utility of the various consequences. Notice that consistency is a procedural, not a substantive, criterion.

Exclusive preoccupation with outcomes is a serious limitation of decisionism, since social processes seldom have only instrumental value for the people who engage in them. In most areas of social activity, the processes and rules that constitute the enterprise and define the roles of its participants matter quite apart from any identifiable “end state” that is ultimately produced. Indeed in many cases it is the process itself that matters most to those who take part in it.19

Thus, as John Dewey once observed, the most important thing about popular voting and majority rule is less the actual outcome of the voter choice than the fact that the electoral process compels prior recourse to methods of discussion, consultation, and persuasion, and the resulting modification of views to accommodate the opinion of the minority.20

Again, knowledge of the outcomes, even when they can be measured precisely, is not the type of information that citizens and policymakers find most useful in many situations. Simply knowing that outcomes are good or bad, without knowing the process that has produced them, does not tell decision makers and critics very much about what to do. As I will show in chapter 8, knowledge of process is often essential for purposes of evaluation and learning since it provides information that outcome measures are almost sure to miss.

The usefulness of the decisionist approach to public policymaking is further limited by the fact that in politics, as in the law (but not the market), decisions must always be justified. Justificatory arguments play an important role in the policy debate but are alien to the philosophy of decisionism. In part this is because the reasons given to justify or explain a decision are often different from the decision maker’s original motives or “revealed preferences” and thus appear to be mere rationalizations. But as I will show in greater detail in chapter 2, it is not necessarily dishonest or merely “rationalizing” to use arguments based on considerations different from those that led to the adoption of a certain position. There is no unique way to construct an argument: data and evidence can be selected in a wide variety of ways from available information, and there are several alternative methods for analysis and ways of ordering values. There is nothing intrinsically reprehensible in selecting the particular combination of data, facts, values, and analytic methods that seems to be most appropriate to convince the people who have to accept or carry out the decision. If analysts are reluctant or unable to provide such postdecision justifications or explanations, policymakers have no choice but to turn elsewhere for assistance.

Because decisionism is an internally consistent doctrine, the various limitations noted above—the assumption of the unitary decision maker, the lack of concern for process, the failure to distinguish among different classes of decisions and to recognize the role of argument and persuasion in decision making—are all intimately related. Their cumulative effect is to produce an overintellectualized version of policy analysis which gives undue emphasis to the more technical aspects of a subject that in fact should be concerned with the whole of the policy process.

Decisionists look upon policy problems as if they were puzzles
for which, given clear goals and sufficient information, correct solutions always exist and can be found by calculation rather than by the exercise of political skills. Hence, policy-making can be intelligent or rational only if it is preceded by systematic analysis of the alternatives in all their implications. To act rationally is, according to this view, always to do two things: to work out a plan of action and to put into practice what the plan prescribes. It is, in Gilbert Ryle's phrase, to do a bit of theory and then to do a bit of practice.\(^{21}\)

But it is notoriously possible to plan well and to implement the plan stupidly. Moreover, by the original assumption, in order to be rational the planning process itself would have to be preceded by yet another process of planning to plan. This infinite regress reduces to absurdity the notion that for a decision or policy to be intelligent it must be guided by a prior intellectual operation. "Intelligent" cannot be defined in terms of "intellec­tual" or "knowing how" in terms of "knowing that."\(^{22}\)

Someone without a knowledge of medicine can hardly be a good surgeon, but excellence at surgery is not the same thing as knowledge of medical science, nor is it a direct result of it. Like surgery, the making of policy and the giving of policy advice are exercises of skills, and we do not judge skillful performance by the amount of information stored in the head of the performer or by the amount of formal planning. Rather, we judge it by criteria like good timing and attention to details; by the capacity to recognize the limits of the possible, to use limitations creatively, and to learn from one's mistakes; by the ability not only to show what should be done, but to persuade people to do what they know should be done.

Perhaps the most serious limitation of the decisionist view is not that it is wrong per se, but that it has led to a serious imbalance in the way we think about policy-making. The following chapters will attempt to provide a more realistic view of the uses of knowledge and analysis in policy deliberation and a better appreciation of the skills needed to transform ideas into actions.

---

22. Ibid., 32.
idence, to keep many threads in hand, to draw for an argument from many disparate sources, to communicate effectively. He recognizes that to say anything of importance in public policy requires value judgments, which must be explained and justified, and is willing to apply his skills to any topic relevant to public discussion.

The image of the analyst as problem solver is misleading because the conclusions of policy analysis seldom can be rigorously proved. Demonstrative proof that a particular alternative ought to be chosen in a particular situation is possible only if the context of the policy problem is artificially restricted. One must assume that there is no disagreement about the appropriate formulation of the problem, no conflict of values and interests, and that the solution is, somehow, self-executing. Also, the analyst should have all the relevant information, including full knowledge of present and future preferences and of all consequences of all possible alternatives.

The impossibility of proving what the correct action is in most practical situations weakens the credibility of analysis as problem solving, but it does not imply that information, discussion, and argument are irrelevant. We reason even when we do not calculate—in setting norms and formulating problems, in presenting evidence for or against a proposal, in offering or rejecting criticism. In all these cases we do not demonstrate, but argue.

Argumentation differs from formal demonstration in three important respects. First, demonstration is possible only within a formalized system of axioms and rules of inference. Argumentation does not start from axioms but from opinions, values, or contestable viewpoints; it makes use of logical inferences but is not exhausted in deductive systems of formal statements. Second, a demonstration is designed to convince anybody who has the requisite technical knowledge, while argumentation is always directed to a particular audience and attempts to elicit or increase the adherence of the members of the audience to the theses that are presented for their consent. Finally, argumentation does not aim at gaining purely intellectual agreement but at inciting action, or at least at creating a disposition to act at the appropriate moment.

It will be noticed that the distinctive features of argumentation are precisely those which characterize dialectic and rhetorical reasoning. Thus, to recognize that policy analysis has less to do with proof and computation than with the process of argument is to make contact with an old philosophical tradition that defines rationality not in instrumental terms, but as the ability to provide acceptable reasons for one’s choices and actions. By restricting the role of reason to discovering appropriate means to given ends, instrumental rationality relegates values, criteria, judgments, and opinions to the domain of the irrational or the purely subjective. Analysis-as-argument holds that this narrowing of discourse goes against the grain of a system of government by discussion. In order to influence public deliberation in significant ways, analysts must open themselves to a wider range of argument than is allowed by the methodology of decisionism.

It is true that practicing policy analysts often engage in argumentative discourse: they debate values, question objectives, agree or disagree about assumptions, and advocate or justify courses of action on the basis of less-than-conclusive evidence. What is problematic about these practices is not their content but the fact that they remain unexamined and that in consequence crucial aspects of analysis escape critical evaluation. In this chapter I discuss some of the most significant rhetorical uses of policy analysis.

NORM SETTING

It is widely assumed that public deliberation and public policy are primarily concerned with setting goals and finding the means to achieve them. Actually, the most important function both of public deliberation and of policy-making is defining the norms

that determine when certain conditions are to be regarded as policy problems. Objective conditions are seldom so compelling and so unambiguous that they set the policy agenda or dictate the appropriate conceptualization. In the 1950s the issue of poverty was a minor one in American public consciousness. In the 1960s, with little change in the distribution of income, it became a significant part of public policies.

What had changed were attitudes and views on poverty, and beliefs in the capacity of government to find solutions to social problems. A particularly important new element was the emergence of an intellectual consensus about the "structural" causes of poverty. As Charles Murray writes, "The emergence of the structural view of the poverty problem was unexpected and rapid. At the beginning of 1962, no one was talking about poverty; by the end of 1963 it was the hottest domestic policy topic other than civil rights. But it was not just 'poverty' that was being talked about. 'Structural poverty' was now the issue." As was pointed out above, in the decisionist view rational policy analysis can begin only after the relevant values have been authoritatively determined. In fact, these values are neither given nor constant, but are themselves a function of the policymaking process that they are supposed to guide. Of the problems with which a democratic government is expected to be concerned today, many were not regarded as policy problems a century or even a few decades ago. Of those that were so regarded (such as the relief of extreme poverty), the norms have radically changed. Yet the process that has brought about these changes in norms is the same historical process which these norms have guided.

Far from waiting passively for the stipulation of public values to be served, policy analysts and researchers are often deeply involved in the process of norm setting. President Lyndon Johnson's "war" on poverty is one example. Another example is the policy innovation represented by pollution-control laws with clear goals and timetables to achieve them, such as the 1970 Clean Air Act and the 1972 Federal Water Pollution Control Act. This legislation was significantly influenced by a theory of "agency capture," according to which vague statutory language was a cause of the capture of regulatory agencies by business. The proposed remedy was statutes that have clear goals, set fixed deadlines for achieving them, and empower citizen groups to take slow-moving agencies to court.

These ideas were incorporated in a number of influential textbooks and were eventually adopted by Congress in the popularized version provided by members of the Ralph Nader organization and other policy advocates. The final result of the combined efforts of researchers and activists was a radical resetting of norms relating to environmental and health protection. Judged by the new norms the traditional regulatory structure—based on informal negotiation with industry, weak enforcement by state agencies, and a large measure of administrative discretion—suddenly appeared inadequate and prone to corruption. A major shift from decentralized regulation and voluntary compliance toward regulation at the national level by means of legally enforceable standards was the legislative response to the new norms.

Conceptually, norm setting can be distinguished from norm using—the search for solutions that satisfy current norms. This distinction is analogous to the traditional dichotomy of policy and administration. The policy/administration dichotomy has been used to support the doctrine that political leaders make policy while the task of administrators and experts is to find the appropriate means to implement it. But it is not the case that policy settles everything down to a certain point while administration deals with everything below that point. Policy and

administration do not occupy two separate spheres of action, but interact throughout the entire process of policy-making.

One reason why it is difficult in practice to separate policy from administration, or norm setting from norm using, is that legislative mandates are often so vague, ambiguous, or contradictory that there are no clear standards for administrators and experts to apply. Even when the statutes attempt to define goals with great precision, as in the case of the environmental legislation of the early 1970s, available technical and scientific knowledge may be insufficient to indicate ways that unambiguously fulfill the original goals. Because uncertainty is so pervasive in policy-making, the values of administrators and experts inevitably count a great deal.

Hence, in drawing the conceptual distinction between norm setting or policy-making on the one hand, and norm using or administration on the other, we must be careful to avoid any implication that policy and administration occupy two completely separate spheres, or are the responsibility of two completely separate groups of people. Norm setting is not the prerogative of high-level policymakers, nor do administrators and experts deal only with means. In fact, as Charles W. Anderson writes,

the actual role of policy professionalism in contemporary government is probably more prescriptive than instrumental. The setting of standards of good practice is a large part of what professionalism means. Most policy professions are such precisely because they provide standards for public policy. In such diverse fields as forestry, public health, nutrition and welfare, the essential function of the expert is often that of setting criteria for the definition of public objectives and the appraisal of public programs. 7

Experts may play an important role in setting standards for public policy even when they appear to be dealing with purely factual questions. The following example is typical of a wide range of situations in regulatory decision making. 8 In 1974 an environmental group, the Environmental Defense Fund, petitioned the Environmental Protection Agency (EPA) to suspend and cancel two chemical pesticides, Aldrin and Dieldrin (A/D). During the cancellation hearings it became clear that there was no agreement over the standards for inferring carcinogenicity. The experts for Shell Chemical Company, the producer of A/D, argued that certain strict criteria had to be satisfied before a substance could be considered to be carcinogenic. The standards advocated by these experts included traditional toxicologic criteria such as the development of tumors in two or more animal species exposed to the substance in the laboratory, proof that the tumors are substance-related, and the availability of data proving the existence of at least one human cancer. EPA's case against A/D rested on different criteria of carcinogenicity. According to the agency's experts, a carcinogen is any agent that increases the induction of even benign tumors in people or animals; a carcinogenic agent may be identified through analysis of experiments on animals or on the basis of properly conducted epidemiological studies; and any substance that produces tumors in one animal species in appropriately conducted tests must be considered a carcinogenic hazard to man.

Neither set of criteria could be dismissed as being unreasonable or contrary to the rules of scientific evidence. Consequently, the choice had to be made on nonscientific grounds. In objecting to Shell's criterion of at least one A/D-induced human cancer, the EPA's experts maintained that since animal tests were sufficient to predict carcinogenic risk, it was ethically unjustifiable to wait for the demonstration of human harm. They also argued that the dictates of prudent policy implied that positive evidence of tumors in one animal species should supersede negative results in other species.

In advocating standards of evidence that departed significantly from more traditional toxicologic criteria, the EPA's ex-


perts were effectively proposing new norms for public policy concerning carcinogenic risk, including criteria of what constitutes sufficient evidence for public decisions. The general lesson suggested by this and previous examples is clear. Experts, including policy analysts, are often engaged in setting norms rather than in searching for solutions that satisfy given norms. Empirical methods have no point of attack until there is agreement on norms, since the nature of the problem depends on which norms are adopted. Hence, argument and persuasion play the key role in norm setting and problem definition.

Similar conclusions hold in the case of policy evaluation, as I will show in more detail in chapter 8. The prominence achieved by evaluation research in recent years shows that policy analysts have finally come to realize that the effective delivery of public services requires more than the discovery and installation of some theoretically optimal program. Even more important is to learn how the program actually behaves, whether it is accomplishing what was intended, and if not, how it can be improved or discontinued.

It is widely assumed that these are purely empirical determinations, involving neither value choices nor preconceived opinions. In fact, values and opinions count for a great deal in evaluation partly because the outcomes of practice are intrinsically ambiguous under normal circumstances. Data are often poor and the measurement instruments unreliable; in addition, the causes of both success and failure are multiple, and different stakeholders are usually interested in singling out some particular subset of causes. Hence, the assessment of a particular outcome depends on which assumptions and criteria the evaluator adopts. The profound equivocality of praxis, to use Donald Campbell's phrase, cannot be removed by improved measurement and testing techniques, but can be represented and clarified in argumentation and reciprocal persuasion.


When analyzing policy decisions, or decision processes in general, it is useful to distinguish between the procedure by which a conclusion is reached—the process of discovery—and the procedure by which the conclusion may be justified—the process of justification. The way in which a conclusion was reached does not always answer the question of whether the conclusion is in fact reasonable or justifiable. For example, the personal motives that guided the decision maker may be inadequate to explain his decision to others or to persuade them to implement it. If we term those considerations on which a person acts motives and those which may be used in interpersonal communication reasons, then we may say that not all motives need be reasons and that not all reasons function as motives.10

Similarly, the way a solution to a mathematical or scientific problem is discovered is not always or even usually the way in which the solution is presented, justified, or defended to the community of specialists. Even in scientific problem solving the private moment of intuition must be followed by a public process of justification and persuasion.

The distinction between motives and reasons, or between discovery and justification, is not sufficiently appreciated by analysts and other students of policy-making, but its significance has been clearly recognized by legal scholars and philosophers of science. Consider the case of a judge who decides a case on the basis of his subjective notion of fairness, a hunch that a particular decision would be right, while realizing at the same time that considerations of this kind do not count as justifications for a binding determination. Thus, the judge frames his opinion in the objective categories of legal argument, and any subsequent developments in the case (for example, an appeal) will be based on the published opinion, not on the actual process followed by the judge in coming to the conclusion. It is a fact of great methodological interest that most legal systems allow the opinion stat-

ing the reasons for a judicial decision to follow rather than precede that decision. Also, different judges may agree on a decision but disagree about the best way to justify it; in the American system they are given the opportunity to present their positions in separate arguments.

Such procedural rules must seem absurd to somebody who assumes that a judicial opinion is an accurate description of the decision process followed by the judge in coming to a conclusion. If, however, the opinion is viewed as a report of justificatory procedures employed by the judge, then the appeal to legal and logical considerations, which possibly played no role in the actual decision process, becomes quite understandable. In fact, the judge's opinion is not the premise of a syllogism that concludes in the decision; it is, rather, a means of increasing the persuasive force of the decision and exercising rational control over conclusions that may be suggested by extralegal considerations.

Contrary to what positivism would have us believe, justificatory arguments also play an important role in science. Scientific arguments, it is now recognized, are attempts to make some theory extremely plausible and convincing, but cannot be conclusively proved either by mathematics or by inductive procedures. In the words of physicist John Ziman, scientific reports “are not diaries or journals, telling us exactly what occurred in a particular laboratory on a particular day. They give, rather, a carefully edited version of such events, and inform us what ought to happen if you try to repeat the experiment yourself under the prescribed conditions.” The communication of the experimenter to his colleagues is not merely an exposition of what happened when certain operations were performed; rather, it is an attempt to convince them that the world behaves as the scientist has conceived it. After the private moment of discovery, “there must come the public demonstration, the deliberate process of persuasion. That is why I say that a good experiment is a powerful piece of rhetoric; it has the ability to persuade the most obdurate and skeptical mind to accept a new idea; it makes a positive contribution to public knowledge.”

Justificatory arguments play an even larger role in policymaking. To decide, even to decide correctly, is never enough in politics. Decisions must be legitimated, accepted, and carried out. After the moment of choice comes the process of justification, explanation, and persuasion. Also, policymakers often act in accordance with pressures from external events or the force of personal convictions. In such cases arguments are needed after the decision is made to provide a conceptual basis for it, to show that it fits into the framework of existing policy, to increase assent, to discover new implications, and to anticipate or answer criticism.

Moreover, since policies exist for some time, their political support must be continuously renewed and new arguments are constantly needed to give the different policy components the greatest possible internal coherence and the closest fit to an ever-changing environment. Policy development does not consist of taking first this decision and then that, piecemeal. Rather, as I will argue in chapter 7, the process of policy development is guided by a parallel intellectual process of refining and developing some original policy idea.

Thus, postdecision arguments are indispensable in policymaking, yet they have been traditionally dismissed as attempts at “rationalization.” Indeed, one of the recurring criticisms of analysis is that it provides “pseudo-scientific rationalizations” for politically or bureaucratically determined positions. For example, it was said that Robert McNamara, the former U.S. secretary of defense, used “studies showing greater cost-effectiveness for passive rather than active defense . . . as ammunition against congressmen who wanted [the antiballistic missile], even though McNamara’s opposition to ABM was based on other factors.”

Whether or not this particular allegation is true, it is not necessarily dishonest or merely “rationalizing” to use arguments

13. Ibid., 36.
based on considerations different from those that led to the adoption of a certain position. We have already emphasized the fact that arguments are not formal proofs. A logical or mathematical proof is either true or false; if it is true, then it automatically wins the assent of any person able to understand it. Arguments are only more or less plausible, more or less convincing to a particular audience. It has also been pointed out that there is no unique way to construct an argument: data and evidence can be selected in a wide variety of ways from the available information, and there are several alternative methods of analysis and ways of ordering values. Hence, there is nothing intrinsically reprehensible in selecting the particular combination of facts, values, and methods that seems to be most appropriate to convince a particular audience.

The importance of postdecision arguments for rationalizing actions and guiding policy development is particularly evident in economic policy-making. For example, President Franklin D. Roosevelt's policy of increased government spending to reduce unemployment and get out of the depression has been called Keynesian. But Roosevelt did not have to learn about government spending from Keynes. The idea that the influence of the British economist lay behind the policies of the New Deal began to take root fairly early, but it is only a legend.\(^{15}\) The theories of Keynes merely provided a sophisticated rationale for what Roosevelt was doing anyway. The answers that these theories provided to questions about the causes of long-term unemployment and the reasons for the effectiveness of public spending were not prerequisites for Roosevelt's expansionist fiscal policy. But as these answers came to dominate the thinking of economists and politicians, they helped to make expansionist fiscal policy the core idea of liberal economic policy for several decades. In the words of Herbert Stein, a former chairman of the President's Council of Economic Advisers, "Without Keynes, and especially without the interpretation of Keynes by his followers, expansionist fiscal policy might have remained an occasional emergency measure and not become a way of life.\(^{16}\)

Thus, it is wrong to assume that the only legitimate use of analysis is to assist the policymaker in discovering a solution to a problem. Policymakers need retrospective (postdecision) analysis at least as much as they need prospective (or predecision) analysis, and probably more. That this kind of analysis is shunned by many analysts can only be explained by the constraining hold of the decisionist methodology on their minds. As I argued in chapter 1, a serious limitation of this methodology is precisely its failure to appreciate the significance of the rhetorical aspects of policy-making—the role of justification, communication, and persuasion in the formation and development of public policy.

The fundamental reason for failing to appreciate these aspects of policy-making is now clear. As long as rationality is defined as choosing the best means to a given end, it is natural to consider retrospective justificatory arguments as being outside the pale of professional analysis—"mere rhetoric," propaganda, or rationalization. However, this instrumental view is not an adequate characterization of the role of reason in human affairs. The social psychologist Karl E. Weick has made the point with particular cogency:

> Rationality makes sense of what has been, not what will be. It is a process of justification in which past deeds are made to appear sensible to the actor himself and to those other persons to whom he feels accountable. It is difficult for a person to be rational if he does not know precisely what it is that he must be rational about. He can create rationality only when he has available some set of actions which can be viewed in several ways. It is possible for actors to make elaborate, detailed statements of their plans. However, the error comes if we assume that these plans then control their behavior. If we watch closely, it will become clear that the behavior is under the control of more determinants than just the vocally stated plan. And at the conclusion of the actions, it will never be true that the plan as first stated will have been


exactly accomplished. But something will have been accomplished, and it is this something, and the making sense of this something, that constitute rationality.

Thus, modern psychological theory reinstates the classic notion of rationality: an action is rational if it can be explained and defended by arguments acceptable to a reasonable audience. In this old-new perspective any sharp distinction between discovery and justification, or between reasons and rationalizations, appears artificial and unrealistic.

ADVOCACY

Equally artificial and difficult to sustain in practice is the related distinction between policy analysis and policy advocacy—between laying out the alternatives that can accomplish a given goal and advocating changes in what governments do. Analysts with extensive experience in advising policymakers in business and government point out that clients want and need advice about objectives as well as the most efficient ways of achieving them.

Economist Carl Kaysen goes as far as arguing that in his role as adviser, the economist "functions primarily as a propagandist of values, not as a technician supplying data for the pre-existing preferences of the policy makers...The adviser becomes, in fact, a supplier of arguments and briefs which seek to gain wider support for economists' political values..." This is an extreme opinion with which few economists or other social scientists would agree, but it is the understandable reaction of a practicing analyst caught between the impossible demands of an outdated methodology and a widespread unwillingness to discuss openly the rhetorical aspects of his craft.

The prevailing positivistic methodology of the social sciences stresses the separation of facts from values and prizes objectivity and a willingness to report findings whether or not they agree with one's preferences or expectations. But as we said earlier, values and preferences are shaped by experiences; the choice of means helps to alter the criteria by which the correctness of the means must be judged. Although people consider what to do before they act, they act in the light of what they are already doing and of what is presently happening. The analyst cannot bring the policy process to a stop while goals are defined and values clarified, and then set everything once more in motion.

Moreover, the findings of social science are usually open to a variety of explanations and interpretations. Like the issues of regulatory science mentioned in chapter 1, many questions investigated by social scientists are trans-scientific in the sense that they can be stated in the language of science but cannot be answered in strictly scientific terms. Few theories advanced by social scientists can be tested by means of controlled experiments. At any rate, it would be too expensive and time-consuming to generate the kinds of data required to refute the various theories proposed on such issues as the causes and possible remedies of crime and illiteracy, the relation between education and earnings, or that between employment and inflation. Hence, any particular set of facts will be consistent with a variety of theories and hypotheses. Since the official methodology provides no objective criterion for choosing under these circumstances, analysts cannot be blamed for selecting the explanation that best fits their preconceived opinions or expectations. The fault lies not in using subjective criteria but in leaving those criteria unexamined.

In addition, the job of the analyst is not only to find solutions within given constraints, but also to push out the boundaries of the possible in public policy. Major policy breakthroughs become possible only after public opinion has been persuaded to accept new ideas. But new ideas face powerful intellectual and institutional obstacles. Economic, bureaucratic, and political interests combine to restrict the range of options that are submitted to

19. Aaron, Politics and the Professor, 164–67.
ANALYSIS AS ARGUMENT

...public deliberation or given serious consideration by the experts. Because of intellectual and institutional inertia, ideas in agreement with current practices and accepted doctrine usually enjoy a considerable comparative advantage over unconventional proposals. At the same time, new ideas generally lack strong empirical and theoretical support. Time is needed until favorable evidence accumulates and auxiliary theories come to the rescue. For all these reasons, objective analysis, unassisted by advocacy and persuasion, is seldom sufficient to achieve major policy innovations.

Thus, in order to be effective, an analyst must often be an advocate as well. But he is also a firm believer in the virtues of the scientific method, and this belief is generally associated with a distaste for advocacy and persuasion. One way to defuse the conflict between practical effectiveness and scientific integrity is to note that many outstanding scientists have not been loath to use persuasion when the situation seemed to require it. For example, eminent historians of science like Duhem and Koyré have likened the work of Galileo to propaganda. "But propaganda of this kind is not a marginal affair that may or may not be added to allegedly more substantial means of defense, and that should perhaps be avoided by the 'professionally honest scientist.' In the circumstances we are considering now, propaganda is of the essence. It is of the essence because interest must be created at a time when the usual methodological prescriptions have no point of attack; and because this interest must be maintained, perhaps for centuries, until new reasons arrive."20

As one would expect, the role of persuasion is even more significant in the social sciences. Thus, in discussing Adam Smith's principles of division of labor and free exchange, the authors of a well-known textbook on economics write: "It is interesting that Smith's book did not contain a logically correct exposition; instead it contained a masterfully persuasive statement of the results of free exchange. It was Robert Torrens who, some forty years after the idea had been 'sold,' demonstrated its logical validity. Possibly, had Smith tried to give a logically air-


tight demonstration, instead of a suggestive, plausible interpretation, he would never have made his 'point' popular."22 George Stigler adds Jevons and Böhm-Bawerk to the list of outstanding economists who "have employed the techniques of the huckster." According to Stigler, persuasive arguments have preceded and accompanied the adoption on a large scale of almost every idea in economic theory.

If advocacy and persuasion play such an important role in the development of scientific ideas, can policy analysts afford to slight them in the name of a historically mistaken view of scientific method? In policy analysis, as in science and in everyday reasoning, few arguments are purely rational or purely persuasive. A careful blend of reason and persuasion is usually more effective than exclusive reliance on one or the other. Style, elegance of expression, and novel modes of communication are often important means of winning support for a new idea and overcoming preconceived hostility and institutional inertia. The practical question, therefore, is not whether to use persuasion, but which form of persuasion to use and when. There are in fact, as we shall see in a moment, situations where the use of persuasion, far from violating the analyst's code of professional behavior, is not only effective but also rationally and ethically justifiable.

ADVICE AND PERSUASION

To examine in more detail the role of persuasion in analysis, let us consider the important special case of policy advice. Advice is sought and given in different situations depending, among other things, on the clarity of the objectives of the policymaker, his understanding of the problem situation, and the knowledge

and stance of the adviser. Three situations deserve to be singled out for special attention.23

In the first situation, the task of the adviser is to determine the best, cheapest, or most effective way of achieving an objective that the policymaker has already decided to aim at. Here the advice takes the form “If you want A, then do B,” as in the economist’s prescription “If you want to maximize profits, set production at the level where marginal revenue equals marginal cost.” The assumption implicit in such prescriptions is that the problem has a definite solution and that there exists a well-defined procedure which, if followed, will enable the advisee to achieve his objectives. In other words, the advisee’s goal is clear and the path to it is also clear, though not to the advisee; the analyst only needs to work out the answer. Recommendations made in situations of this type are better described as instructions or prescriptions than as advice.

We meet a somewhat different situation when the policymaker’s problem involves more than selection of the most appropriate means to achieve a given end. For example, the policymaker may be uncertain about the nature of the problem to be solved; he feels that things are not as they should be, but has no clear idea about what should be done. Even when the problem situation is reasonably well defined there may be several alternative formulations or methods of solution but no standard way of choosing among them. In such cases, which are quite frequent in practice, it is appropriate to say that the analyst gives advice rather than instructions or prescriptions, as in the first case.

Finally, we may identify a third type of situation in which the analyst uses the language of advice to redirect the policymaker’s attitudes, preferences, or cognitive beliefs. If, for example, the analyst feels that the policymaker has incorrectly formulated the problem, he may feel obliged to persuade him to accept his (the analyst’s) formulation. In such cases, one should perhaps speak of persuasive advice; the question is, under what circumstances


is this form of advice rationally and morally justifiable? Notice that when the analyst uses persuasion, he is always acting, at least in part, as advocate rather than disinterested adviser.

Sometimes persuasion is a necessary preliminary to get the policymaker’s or the public’s attention, to make them “listen to reason” when they are blinded by stereotypes or by wishful thinking. Walter Heller, a former chairman of the Council of Economic Advisers, gives this example: “In 1961, with over five million unemployed and a production gap of nearly $50 billion, the problem of the economic adviser was not what to say, but how to get people to listen. Even the President could not adopt modern economic advice, however golden, as long as the Congress and the public ‘knew’ that it was only fool’s gold.... Men’s minds had to be conditioned to accept new thinking, new symbols, and new and broader concepts of the public interest.”24 As Heller suggests, policymakers tend to think in traditional categories, or in terms of alternatives that are unduly restricted in relation to their own objectives. Persuasion is needed to induce them to consider different formulations or approaches to the issue under discussion since the psychological effect of factual arguments may not be strong enough to overcome the inertia of long-established patterns of thought. Experience also shows that facts and statistics are seldom sufficient to bring about changes in behavior—even after the need for a change has become clear. For this reason, public policies to induce citizens to adopt healthier lifestyles tend to rely at least as much on persuasion as on objective information.

Next consider the case where the motivation to attack a persistent problem, such as crime or illiteracy, is in advance of the knowledge required to solve it. It may be that the technical tools for an adequate treatment of the problem do not yet exist, or that good evidence on causal factors is hard to get. In such cases popular appeals and persuasion, bolstered by whatever empirical and theoretical knowledge is available, may succeed in stimulating interest in the issue and keeping it alive until satisfactory methods of solution have been developed.

When knowledge of a problem is very limited, experts tend to disagree about its causes and possible solutions and are thus unable to provide unequivocal advice. In this situation it is usually better to let each expert openly advocate his position, preferably in a well-structured adversary setting, than to attempt to enforce a consensus in the name of scientific objectivity. Adversary procedures are specifically designed to bring out unstated assumptions, differing interpretations of the facts, and gaps in logic or in the evidence. Thus, they provide powerful incentives for the adversaries to present the strongest arguments in favor of their respective positions.

Because of these advantages, the adversary process, in a form called "multiple advocacy," has been recommended as a way of organizing expert advice in areas of public policy such as national security and economic policy-making at the presidential level. The basic assumption underlying multiple advocacy is that a competition of ideas and viewpoints, rather than reliance on analyses and recommendations from advisers who share the perspective of the policymaker, is the best method of developing policy. Multiple advocacy is a process of debate and persuasion designed to expose the policymaker systematically to competing arguments made by the advocates themselves. Through the offices of an "honest broker" it attempts to ensure that all interested parties are represented in genuinely adversarial roles, and that the debate is structured and balanced. 25

In all the cases we have discussed, persuasion is justifiable on professional as well as ethical grounds. It is also important to keep in mind that since policy analysis cannot produce conclusive proofs but only more or less convincing arguments, persuasion always has a role to play in increasing both the acceptability of advice and the willingness to act on less than complete evidence. For this reason, experienced analysts suggest that analysis should be done in two stages: the first stage to find out what the analyst wants to recommend, and a second stage to make the recommendation convincing even to a hostile and disbelieving audience. 26 This is sound advice as long as it is not interpreted to imply that communication and persuasion are discrete and separable parts of analysis rather than pervasive aspects of the analytic process.

Throughout this chapter I have emphasized that in order to be persuasive, evidence and arguments must be chosen with a particular audience in mind: the same conclusions may have to be justified differently in different contexts.

As Arnold Meltsner points out in an interesting essay on communication in policy analysis, concentrating on the immediate client as the sole recipient of advice and information may be dangerous. 27 In a complex organization or political system, the immediate client is only one of numerous actors who comprise the analyst's audience, and the analyst may be mistaken in focusing his communication on that single client. Too, it is often difficult to prevent the dissemination of analytic information through press reports, public debate, or deliberate leaks. In principle, any reader of a policy study may be considered a member of the analyst's audience. Finally, it is possible that by the time the analysis is completed, the original client will have been replaced, key elected and appointed officials will have left office, and other actors will have moved on to another problem.

For all these reasons, audience—a term with a long tradition in rhetoric—is a better, more flexible, and more neutral characterization of the set of actual or potential recipients and users of analysis than more familiar terms like client or decision maker. It also reminds us that the main justification of advocacy and persuasion in democratic policy-making is their function in a continuous process of mutual learning through discourse.


The argumentative model of analysis assumes that analysts can seldom demonstrate the correctness of their conclusions, but only produce more or less persuasive evidence and reasonable arguments. However, belief in the possibility of discovering correct solutions for a wide variety of problems has been important historically in legitimating the use of analysis in policy-making. Hence, the admission that analysis is fallible raises professional as well as political questions. First, is it at all possible to define standards of quality for policy analysis if its conclusions are always tentative and open to refutation? Or, to put the same question in a slightly different form, how can analysts steer a safe course between the Scylla of absolute certainty and the Charybdis of methodological anarchism—the easy philosophy of "anything goes"? Finally, how can policy analysis be justified and legitimated once the claim to certainty of conclusions is abandoned?

This chapter provides some answers to these questions by looking at the craft aspects of analysis, the details of the process by which policy arguments are produced. My approach reflects recent developments in the philosophy and sociology of science.

Few scientists and philosophers of science still believe that scientific knowledge is, or can be, proven knowledge. If there is one point on which all schools of thought agree today, it is that scientific knowledge is always tentative and open to refutation. And while the older history of science was little more than a chronicle of the irresistible advance of the different sciences, the contemporary historian tries to understand "how such sciences can succeed in fulfilling their actual explanatory missions, despite the fact that, at any chosen moment in time, their intellectual contents are marked by logical gaps, incoherences, and contradictions." 1 Even mathematical knowledge is fallible, tentative, and evolving, as is every other kind of human knowledge. 2

Since the endeavors of individual scientists are fallible, the emergence of a (provisionally) accepted body of knowledge must be explained in terms of social mechanisms of evaluation and quality control. "Nature," writes Jerome Ravetz, "is not so obliging as ever to give marks of True and False for scientific work, and so a scientific community sets its standards for itself." 3

As contemporary epistemologists emphasize, some form of conventionalism is the inescapable logical consequence of fallibilism. If there is no demonstrative certainty for the conclusions of science, their "truth," or at any rate their acceptability as scientific results, can only be established by convention: through a consensus of experts in the field and the fulfillment of certain methodological and professional norms—the rules of the scientific game.

In planning an experiment, evaluating a batch of data, or choosing among alternative research strategies, the scientist utilizes knowledge and skills that are not themselves scientific but are acquired by practice and imitation. In the language of chapter 1, the process of scientific research (as distinct from the fin-

ished products of research) depends more on "knowing how" than "knowing that"; it is a craft, a social process, rather than a purely logical activity.

The craft skills of the scientist form a repertoire of procedures and judgments that are partly personal, partly social and institutional. Thus, in deciding whether a batch of data is of acceptable quality, the scientist applies standards that derive from his own experience but also reflect the professional norms of teachers and colleagues, as well as culturally and institutionally determined criteria of adequacy. Because of the concentration of the older philosophy of science on the logical status of achieved scientific knowledge, the craft aspects of scientific investigation have been neglected until recently. Yet without an appreciation of such aspects "there is no possibility of resolving the paradox of the radical difference between the subjective, intensely personal activity of creative science, and the objective, impersonal knowledge which results from it."

THE ANALYST AS CRAFTSMAN

Craft knowledge—less general and explicit than theoretical knowledge, but not as idiosyncratic as pure intuition—is essential in any kind of disciplined intellectual inquiry or professional activity. It is especially important in policy analysis. The structure of an analytic argument is typically a complex blend of factual propositions, logical deductions, evaluations, and recommendations. Along with mathematical and logical arguments it includes statistical inferences, references to previous studies and to expert opinion, value judgments, and caveats and provisos of different kinds. As we remarked before, this unavoidable complexity rules out the possibility of any formal testing—of proving or refuting the final conclusions. Whatever testing can be done will have to use a variety of criteria derived from craft experience, including the special features of the problem, the quality of the data, the limitations of the available tools, and the requirements of the audience. Only a detailed examination of the different components of the task of the analyst qua craftsman can help the producer or user of analysis steer a reasonable course between unhelpful counsels of perfection and methodological anarchism.

The common-sense notion of craft includes, as basic elements, a body of skills that can be used to produce useful objects; careful attention to the quality of the product; and a sense of responsibility both to the ends of the client and to the values of the guild. In light of these characteristics, to speak of the analyst as a craftsman is actually more than merely using a metaphor: the similarity between the work of the analyst and that of the traditional craftsman is real. In policy analysis, as in the traditional crafts, successful performance depends crucially on an intimate knowledge of materials and tools, and on a highly personal relationship between the agent and his task. Good analytic work cannot be produced mechanically any more than handicraft can be mass-produced. Style can play as big a role in determining the value and acceptability of the analytic product as it does for the results of the craftsman's work.

There are, of course, obvious differences. The craftsman uses concrete materials in order to produce an object that has an appropriate shape and fulfills a specific function. The analyst, on the other hand, operates with concepts, theories, data, and technical tools to produce arguments and evidence in support of certain conclusions. In spite of these differences, craftsmanship is an essential element of any skillful performance. In fact, the classical Aristotelian analysis of craft work has been usefully applied to scientific inquiry5 and will be shown to be relevant also to a discussion of the anatomy of the analyst's task.

Aristotle's scheme involves four constituents (or "causes") of the craftsman's task: material, efficient, formal, and final. They refer, respectively, to the physical substance that is worked on; to the tools the craftsman uses in shaping it; to the form or shape

---

4. Ibid., 75.
6. Ravetz, Scientific Knowledge, chap. 3.
acquired by the substance; and to the purpose of the activity—the creation of an object that fulfills certain functions. To adapt this scheme to the analyst's work, the material component should be identified with the data and information that are used in defining the problem. Tools and techniques are the efficient component of the analyst's task. The "form" of the task is an argument in which evidence is cited and from which a conclusion is drawn, while the final component is the conclusion itself, with the related activities of communication and implementation.

This description of the analyst's task enjoys some advantages over more familiar descriptions. For example, the categories suggested by the decision-making approach (goals, alternatives, consequences, criteria of choice) focus attention on a rather narrow use of analysis in choice-making situations. The craft paradigm, on the other hand, provides categories—data, information, tools, evidence, arguments, conclusions—that are applicable to any type and style of analysis, prospective or retrospective, descriptive or prescriptive, or in an advocacy mode. In turn, these categories clarify important but often overlooked distinctions and, at the same time, facilitate the critical appraisal of the different stages of the analytic process.

DATA, INFORMATION, EVIDENCE

By way of example and as a first application of the craft paradigm, I shall elaborate the distinction already made in chapter 1 among data, information, and evidence. These terms are often used interchangeably with unfortunate consequences not only for the clarity of discourse but, more importantly, for our ability to evaluate the quality of policy arguments.

Data are, so to speak, the raw materials necessary for the investigation of a problem, or perhaps the result of the first working-up of such materials. In policy analysis data are often "found" rather than "manufactured"; that is, they are obtained by unplanned observations (as in the case of a time series) rather than by planned experiments. This fact, as will be shown in the next section dealing with pitfalls, requires craft skills that are rather different from, and in many respects more difficult to acquire than, those needed for the analysis of experimental data.

When data are obtained by sampling, the sampling process may be influenced by the method used, the skill of the sampler, and a host of other factors that could lead to results quite unrepresentative of the general situation. Also, data are usually collected according to categorical descriptions that seldom fit perfectly the purposes of any given inquiry. Even when data are produced by experiments, as in the case of some large-scale social experiments conducted in the United States, there is no guarantee that the best experimental design offers sufficient protection against dangers and pitfalls, of which the "Hawthorne effect"—the fact that people may behave differently when they know they are observed—is only one of the best-known examples.

In sum, since perfection of data is impossible, the standards of acceptance will have to be based on craft judgments of what is good enough for the functions the data perform in a particular problem. Such judgments depend both on internal, disciplinary criteria and on criteria related to the nature of the problem: the standards of adequacy applicable to problems of monetary policy, say, are not necessarily relevant to problems arising in education or welfare. Thus, the simple judgment of soundness of data reveals as in a microcosm all the personal judgments and accumulated institutional experience that go into analytic work.

Before being used in an argument, data usually have to be refined into a more useful and reliable form. This transformation of data involves a new set of craft skills, the application of new tools (often of a statistical or mathematical nature), and the making of a new set of judgments. This new phase of the analyst's work, the production of information, can be illustrated by a variety of examples: the calculation of averages and other statistical indices and parameters, the fitting of a curve to a set of points, the reduction of data by means of some multivariate statistical technique. The operations performed on the original data may be involved or quite simple, but they always represent
a crucial step. Through these operations the raw data have been transformed into a new sort of material, and from this point on the analysis is carried out only in terms of these new entities.

Evidence, as the term is used here, is not the same as data or information. It is, rather, information selected from the available stock and introduced at a specific point in an argument "to persuade the mind that a given factual proposition is true or false." As previously noted, an inappropriate choice of data, their placement at a wrong point in the argument, a style of presentation that is unsuitable for the audience to which the argument is directed—any one of these factors can destroy the effectiveness of information as evidence, regardless of its intrinsic cognitive content. Hence, the criteria for assessing evidence are different from those applicable to "facts." While facts can be evaluated by more or less objective tests, the acceptability of evidence depends on a number of features peculiar to a given situation, such as the nature of the case, the type of audience, the prevailing "rules of evidence," and even the persuasiveness of the analyst.

Similarly, the assessment of the strength and fit of the evidence is considerably more complicated than judgments about the validity and reliability of data. For this reason disputes often arise about the acceptability of the conclusions of policy studies—disputes that cannot be settled either by examination of the data and information or by an appeal to accepted criteria of adequacy. The point that Frederick Mosteller and Daniel P. Moynihan make about the 1966 Report on Equality of Education Opportunity (the so-called Coleman Report) has general relevance: "Professional judgments were made in the survey design and analysis that are open to dispute. This is not always a matter of being right or wrong, but simply of the absence of professional consensus on the points involved."^8

The special significance of evidence is most easily recognized in fields such as law and history, where problems involve both complex arguments and large masses of data, but where the reliability and relevance of the information cannot easily be assessed by standard methods. Thus, jurisprudence we find a highly developed "law of evidence" for the presentation and testing of information offered as evidence in court cases. Judicial notice, facts, documents, physical evidence, testimonial evidence, and hearsay are carefully distinguished and subject to different procedural rules—rules that attempt to balance the value of information against the cost of obtaining it.

Similarly, budding historians are taught to distinguish between records (intentional transmitters of facts) and relics (unpremeditated transmitters of facts), between written and oral testimony, archaeological and linguistic evidence, and so on. Faced with a piece of evidence, the historian asks: Is this object or document genuine? Who is its author or maker? How does the statement compare with other statements on the same point? What do we know independently about the author and his or her credibility?^9

Problems of evidence are discussed at length in several modern classics of historical methodology. Books like E. H. Carr's What Is History?, Marc Bloch's Historian's Craft, and Robin Winks's Historian as Detective are rich in insights whose relevance goes well beyond historical scholarship. They are particularly relevant for policy analysts, who are also constantly faced with problems of evidence. Consider, for example, the problem posed to the analyst by the fact that policy actors will often give different accounts of some crucial event—so different, in fact, that it is almost inconceivable that everyone was perceiving the same event.

Why, then, has the literature of policy analysis practically ignored problems of evidence, except for some recent discussions

---

on the use of statistical evidence in policy-making? One reason has already been alluded to in chapter 1: the intellectualist bias of decisionism, with its emphasis on “knowing that” rather than “knowing how,” has led to a general neglect of the craft aspects of policy analysis, including the skills necessary for a critical evaluation of evidence. In turn, this bias is related to the positivistic tradition in the philosophy of science. Being mainly concerned with the logical and epistemological problems of achieved knowledge, this school has paid very little attention to the actual processes of the production of scientific knowledge.

In this, as in other cases, uncritical acceptance of the “scientific method” (or what is thought to be the scientific method) overlooks some important differences between the natural sciences and policy analysis. For if it is true that neither descriptive nor theoretical natural sciences require highly developed skills in testing evidence, beyond those already involved in producing information, this is for good reasons. In the natural sciences one usually has either a rich and reliable body of information with a relatively simple argument, or a complex theoretical argument needing evidence at only a few points. Such situations are rather exceptional in policy analysis, where one typically deals with large amounts of data of doubtful reliability or relevance. It is likely that the lack of suitable rules of evidence has contributed to the mounting level of dissatisfaction with certain applications of policy analysis.

In particular, large-scale, policy-oriented models have come under heavy attack in recent years. Examples of conceptual, technical, and institutional pitfalls in model construction and utilization have been given in a number of review papers and book-length case studies, leading some critics to conclude that large-scale models are of little use in policy analysis and evaluation. What these critical studies have shown is that the path from model to conclusions is long, involved, and beset by difficulties of all kinds. Only rarely can models provide full answers to policy questions. What they can do, if developed and used with care, is provide evidence which, together with other sources of information, may be used in arguments supporting a certain conclusion or recommendation. A good model is merely one type of evidence among others, not the end of the argument, much less the ultimate authority.

But to be used as evidence, models must satisfy certain procedural requirements that make their assessment possible; they have to be “in proper form.” It is not enough for an effective argument to have a particular form, such as the mathematical garb of an econometric model. After all, a mathematical style of presentation is not incompatible with a “black box” approach—it may even encourage it—and black-box models, it has been rightly said, “will never have an impact on policy other than through mystique, and this will be short-lived and self-defeating.”

A model is in proper form if the assumptions and the evidence used in it are presented in a sequence of steps that conform to basic rules of procedure analogous to those used when questions of law are debated in a court. This is far from being standard practice, as I pointed out in chapter 1. Hence, while it is important to insist, with the critics of large-scale policy models, that models be transparent and as simple as possible, it is also necessary to develop more detailed procedural guidelines, if models are to play their limited but potentially useful role in the policy process. It is indeed ironic that while data generation absorbs so much of the modeler’s time and ingenuity, the transition from...
data to conclusions should often rest on arguments that do not bear close scrutiny. No amount of technical sophistication can compensate for carelessness in structuring the arguments or in drawing the necessary distinctions among data, information, evidence, and conclusions.

**PITFALLS AND FALLACIES**

The craft aspects of policy analysis are revealed most clearly by the concept of *pitfall*. A pitfall is a conceptual error into which, because of its specious plausibility, people frequently and easily fall. It is the taking of a false logical path that may lead the unwary to absurd conclusions. A pitfall is for the practical arguments used in policy analysis what the logical fallacy is in deductive reasoning. In both cases, one has to be always on guard against hidden mistakes that can completely destroy the validity of a conclusion.

Logicians distinguish between a fallacy and a simple falsity. A single statement may be false, but what is fallacious is the transition from a set of premises to a conclusion. Similarly, in policy analysis pitfalls should not be confused with blunders or errors that may affect, for instance, the numerical value of a solution but not the basic structure of the argument supporting it.

In logic there is a long tradition of systematic discussion of fallacies that goes back to Aristotle. Nineteenth-century treatises on logic always included one or more chapters on fallacies. John Stuart Mill devoted Book V of *A System of Logic* to an account and a new classification of fallacies, and A. DeMorgan, while rejecting previous attempts to produce exhaustive descriptions of all possible types of fallacy, nevertheless devoted an entire chapter of his *Formal Logic* to a penetrating analysis of many of the traditionally listed pitfalls.

W. Stanley Jevons, one of the fathers of neoclassical economics, is also the author of a delightful book on elementary logic in which, following Aristotle, fallacies are divided into logical fallacies (those which occur in the mere form of a statement and, thus, could in principle be discovered without any knowledge of the subject matter with which the argument is concerned) and material fallacies (which, being connected with the subject of the argument, can only be detected by those acquainted with the subject).\(^6\)

Among the material fallacies Jevons listed are the fallacy of accident (arguing from a general rule to a special case, where a certain accidental circumstance renders the rule inapplicable, or arguing from a special case to a special case), the irrelevant conclusion (ignoratio elenchi, or arguing to the wrong point), the petitio principi (begging the question), the fallacy of the consequent (non sequitur) and the false cause (post hoc ergo propter hoc, or the assertion that one thing is the cause of another simply because it precedes, or accompanies, it). More recent contributions, mainly devoted to a discussion of fallacies in everyday reasoning, are Robert H. Thouless's *How to Think Straight* and Susan Stebbing's *Thinking to Some Purpose*.\(^7\)

Outside logic and philosophy, the amount of attention devoted to the topic of pitfalls varies considerably among different disciplines. Very few natural sciences have standard literature on the possible pitfalls of their characteristic patterns of argument. This is not surprising, since in the natural sciences it is usually possible to make practical tests of theoretical conclusions, while the existence of effective mechanisms of quality control ensures the rapid elimination of gross fallacies. Also, laboratory courses help the student develop an intuitive feeling for the possibility of pitfalls in the standard procedures by which he verifies theoretical results.\(^8\)

On the other hand, the literature of statistics, a discipline specifically concerned with the logic of inductive reasoning and the weighing of evidence, contains many insightful discussions

---

ANALYSIS AS CRAFT

of pitfalls, both at the technical level—perhaps exemplified at their best by the published discussions of the British Royal Statistical Society—and in textbooks and popular expositions. To some extent, this tradition has been carried over into the neighboring field of econometrics, but it does not seem to have penetrated deeply into actual econometric practice.

An excellent collection of economic pitfalls, which every young policy analyst should read, is E. J. Mishan's Twenty-one Popular Economic Fallacies. But, generally speaking, the social science literature reveals only a marginal awareness of the conceptual significance of pitfalls; for instance, the only detailed discussion of the topic in the eight-volume International Encyclopedia of the Social Sciences is the perceptive article on statistical fallacies by I. J. Good. As a result of this general lack of concern for the ever-present danger of pitfalls, certain types of inference rather widely used in past social science research have only recently been recognized as fallacious.

A good example from the sociological literature is W. S. Robinson's discovery of the so-called ecological fallacy. The pitfall consists of using ecological correlations (that is, statistical correlations involving properties of groups of individuals) as substitutes for individual correlations in which the correlates are properties of individuals: for instance, per capita rates of cigarette smoking and death rates from lung cancer in a number of countries, or race composition and literacy rates in different sections of a country. Correlation coefficients calculated in this way from national or regional averages can be quite misleading, since replacing each country or region by average values eliminates the spread around the averages and thus gives a wrong impression of tight clustering. Robinson has shown that individual and ecological correlations are in general different (they may even differ in sign) and that the values of ecological correlations strongly depend on the type of grouping used. These conclusions cast strong doubts on the validity of a number of empirical studies conducted in the past.

Even the great French sociologist Emile Durkheim has fallen into this pitfall in his famous study of suicide. From data on percentages of literates and suicide rates for each province in nineteenth-century Italy, Durkheim takes averages over three clusters of provinces, obtaining a correlation close to 0.9. From this he concludes that "popular education and suicide are distributed exactly in the same way." Actually, correlation calculated from averages for each individual province turns out to be equal to 0.6, and even this value probably exaggerates the strength of the relationship.

PITFALLS OF ANALYSIS

As already noted, in the natural sciences detailed discussions of different kinds of pitfalls can, to a large extent, be dispensed with because of the stock of practical knowledge that scientists have accumulated from long and successful experience. In policy analysis, however, direct verification of conclusions is seldom possible, while professional mechanisms for controlling the quality of analyses are still in an embryonic stage—the approach is too new for a widely shared tradition of critical thought to have developed.

Also, policy analysts have different disciplinary backgrounds, while students acquiring specialized training in undergraduate and graduate programs of public policy are exposed to academic curricula that vary from school to school and represent tentative

compromises among different intellectual traditions. Most of the technical tools that the budding analyst is now required to learn have been developed by other disciplines, and textbook treatments tend to emphasize only those aspects that appear to have immediate practical usefulness. But concepts and techniques removed from their disciplinary matrix tend to become stereotypes, and their limitations are not easily perceived by people interested only in immediate applications. Such are the roots of some common pitfalls of analysis that B. O. Koopman has labeled "linearitis" (the erroneous belief that everything is linear), "maximitis" (the belief that the only or main purpose of analysis is to maximize something), and "mechanitis" (blind faith in the power of the computer and other mechanical aids). For the same reasons, all the subtlety of statistical reasoning is often lost in ritualistic and almost meaningless applications of hypothesis testing and significance levels.

Hence, a systematic study of pitfalls should become an important part of the training of policy analysts and public managers. As the philosopher of science Jerry Ravetz writes: "A recognition and systematic use of the phenomenon of pitfalls might be very effective in the teaching of those simple but essential craft skills which are involved in scientific, scholarly, or administrative work. An exposition of standard techniques in terms of the pitfalls they are designed to circumvent, with examples, could go far to make them meaningful and obviously worth mastering." In fact, a number of standard works in the early literature of systems and policy analysis include fairly extensive treatment of pitfalls, and some of these discussions have attained the status of minor classics of the discipline. This tradition is continued in a book edited by myself and Edward S. Quade, in which a number of experienced analysts survey the entire spectrum of possible pitfalls of analysis, from problem formulation to implementation.

Here it is not possible to do more than to treat selectively some of the more common types of pitfalls. It will be convenient to organize the discussion under four headings that closely correspond to the four components of the analyst's work which have been identified above: data and information; tools and methods; evidence and argument; and conclusions.

### Data and Information

Analysis usually begins with something less structured than a problem, namely, a problem situation. This is an awareness that things are not as they should be, but without a clear idea of how they might be corrected. Problem setting is the process of translating a problem situation into an actual policy problem stating the goals to be achieved and a strategy for accomplishing them.

The amount of detailed information that is useful at the stage of problem setting is different from what is needed when searching for solutions within a given formulation. The appropriate modes of inquiry will correspondingly differ in the two situations. Because of the unstructured character of a problem situation, imagination, judgment, and analogical and associative thinking play a bigger role in problem setting than rigor and technical skills.

Since the questions asked and the methodological decisions taken at this stage effectively condition the subsequent analysis, one must constantly be aware of the possibility of "preselecting" the final conclusions. For example, it may seem natural for an environmental agency working on a specific pollution problem to consider only alternatives falling within its jurisdiction. Even the competent analysts of the Delaware Estuary Comprehensive Study (DECS), discussed by Bruce Ackermann and his coauthors, chose to define their inquiry to the Delaware Estuary, for which
they were directly responsible, instead of considering other, and probably better, alternatives for outdoor recreation. 28

The next task facing the DECS analysts was to define the nature of the pollution problem in the Delaware River. Traditionally, the amount of dissolved oxygen (DO) in the water has served sanitary engineers as the principal indicator of water quality. It is therefore understandable that the DECS engineers would take recourse to the received wisdom of their profession. But this technical decision had major policy consequences. Once DO was accepted as the relevant indicator, the section of the river experiencing the most serious pollution problem and urgently requiring public action was automatically identified: it was the heavily industrialized region between Philadelphia and Wilmington suffering the most acute oxygen sag. Although there was no reason to believe that increasing DO to feasible levels in this critical section of the river would make the Delaware more suitable for recreational and other uses, this became in fact the focus of the technical analyses. And because the economists responsible for the cost-benefit part of the study accepted uncritically the engineers' formulation of the problem, the policy alternatives from which the agency had to choose came to be expressed in terms of different DO levels. Incidentally, this instructive example reveals another pitfall whose importance has not been sufficiently recognized in the literature.

It is often said that any analyst who accepts the client's initial formulation of the problem uncritically is heading for a disaster. The case of the DECS' shows that it is an equally serious mistake to accept uncritically the initial problem formulations of technical experts. The production and interpretation of data present peculiar problems at the problem-setting stage of the analysis. Thus, information obtained through opinion polls or attitude surveys is particularly sensitive to the way the questions are framed, which in turn may depend on how the problem is conceived. The distinction between facts and artifacts, always a tricky one when social and economic data are involved, becomes particularly difficult when one tries to find out what the problem is. To say, for example, that 7 percent of the labor force is unemployed is only to say that this is the result of applying certain operations embodied in a questionnaire and the answers to it. A simple change in one question, or in some definition, may produce a significant change in the number of unemployed. 29

A related pitfall is the failure to recognize the large margin of error surrounding all socioeconomic statistics. National income statistics probably cannot be known with a smaller margin of error than ± 10 to 15 percent, and comparable uncertainties are present in foreign trade, price, unemployment, and growth statistics. 30 Hence, a reported drop of one or two percentage points in a country's gross national product may or may not be a significant indication that immediate government action is required. Similarly, serious biases in unemployment statistics (due, for example, to various work registration requirements for welfare recipients) have been identified by a number of government bodies and academic experts, but policymakers continue to base their decisions on the existing unreliable statistics. 31

In fact, all social and economic indicators are the product of definition and convention. Measures of inflation, production, education, health, or crime must always be interpreted in relation to a specific context before they acquire any operational meaning; their usefulness depends entirely on a clear recognition of their conventional character. Hence, economic and social statistics cannot be treated in the same way we treat "facts" or physical measurements obtained from a direct apprehension of some natural phenomenon. In Allan Coddington's apt phrase, one cannot, even with good eyesight, go out into the Treasury steps and observe the domestic level of economic activity. Rather, "economic statistics are the result of the bureaucratic compounding

28. Ackermann et al., The Uncertain Search for Environmental Quality.


of enormous quantities of fragmentary and even ambiguous pieces of information: the components are thrown up as a result (or even as a by-product) of the general administrative processes of society, from tax-returns, sample surveys, censuses, and so on; the components are assembled and aggregated by teams of statisticians who have not themselves collected the information. And the amount of information involved is vast. Presumably, we can learn more from imperfect statistics than from no statistics at all, provided that their conventional basis and degree of reliability are appreciated.32

The ad hoc nature of much social and economic accounting—the fact that the classificatory schemes used in data collection are more often dictated by expediency (the availability of data, the feasibility of certain estimates, the standard operating procedures of the data-gathering organization) than by logical or theoretical considerations—is seldom publicized by the data collectors and may be overlooked by the data users. Erich Streissler, a distinguished Austrian economist, tells the story of the econometrician who once made the rather surprising discovery that in Austria profits rose over a certain period exactly as much as wages and salaries. The econometrician had apparently forgotten that in the early days of Austrian national income accounting, profits were assumed—for lack of data—to move in proportion to wages and salaries!33

Sometimes the analyst is in the fortunate position of being able to collect the primary data, or at least of suggesting an appropriate sampling design. Even then the original data are usually much too raw to be used in a model or in an analytic argument without first being refined into a more reliable and useful form. Such refining requires craft skills that are rather different from those used in problem setting and data collection.

The transformation of data into information involves three basic judgments, all of which present the risk of serious mistakes. The first concerns the possible loss of important details: since in most policy problems a "sufficient statistic," containing exactly

33. Streissler, Pitfalls, 28.

the same amount of information as the original sample, does not exist, data reduction may involve too great a loss of information relative to the problem under discussion.

The second is a judgment of the goodness of fit of a model to the original data, a problem for which standard statistical techniques are available. The third basic judgment is that this particular transformation of the data, among the many possibilities, is the significant one.

An entertaining example of a pitfall involving the third kind of judgment is given by Quade: the use, during World War I, of the arithmetic instead of the harmonic mean in computing turnarounds of troop and cargo ships of different speeds.34 This is, admittedly, a rather trivial mistake (though a frequent one, as tests in undergraduate statistics courses show), but precisely its elementary character shows how easy it is to stumble into pitfalls in even the simplest aggregation of data.

Tools and Methods

The tools of policy analysis may be roughly classified according to their function in conceptualization, in data production and manipulation, and in interpretation. The category of conceptual and interpretive tools includes disciplines such as mathematics, economics, and the political and behavioral sciences, which the analyst has to master to some extent in order to do competent work.

I have already alluded to the possibility of pitfalls arising when certain concepts or methods are taken out of their broader disciplinary context. For example, to noneconomists "cost" usually means historical or sunk cost, rather than opportunity cost. A number of early applications of operations research are vitiated precisely by this conceptual mistake. Similarly, "average" is often taken to mean arithmetic or sample average even when, as in the example of the harmonic mean just mentioned, another measure is more appropriate. Again, engineers, being accus-

tomed to operate with input-output ratios, tend to use ratios of benefits to costs as measures of efficiency, disregarding the absolute sizes of benefits and costs.

The danger of conceptual pitfalls is made particularly acute by the prevailing metaphysics, according to which the scientific character of a field is assumed to be in direct proportion to the degree of its mathematical formalization. As a result, the analyst is sometimes tempted to use formal tools that exceed the level of his mathematical or statistical sophistication, and whose range of meaningful applicability he is therefore incapable of assessing.

In disciplines with a long intellectual tradition the introduction of new tools usually opens up lines of research that were previously inaccessible. In newer fields of inquiry, on the other hand, we often witness the phenomenon of "new toolism," a disease to which policy analysts seem to be particularly predisposed. Those affected by this disease "come possessed of and by new tools (various forms of mathematical programming, vast air-battle simulation machine models, queuing models, and the like), and they look earnestly for a problem to which one of these tools might conceivably apply."35

In the preceding pages we have seen how difficult it is to obtain information that is both reliable and relevant. The difficulties are compounded when the data are processed by means of formal techniques and models. For example, are the results derived from a particular model more sensitive to changes in the model and in the methods used to estimate its parameters, or to changes in the data? No general answer to this question seems to be available, and the limited evidence is conflicting. Thus, one econometric study finds that the choice of estimation procedure has more effect on the parameter estimates than the choice of data, while another study concludes that the variations in parameter estimates are generally much greater between different sets of data than between different methods of estimation.36

Regression methods are among the most popular tools of applied social research and policy analysis; yet it is often forgotten that the meaning of a fitted equation

\[ y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_r x_r \]

differs according to whether the x's represent planned or unplanned ("passive") observations. The same formal data manipulations are carried out in both cases, but in the case of passive observations (as in studies correlating income and educational levels, or industrial production and population) it would be quite misleading to interpret the coefficient \( b_i \) as measuring the effect on \( y \) of a unit change in \( x_i \). Such an interpretation is justifiable only if the observations come from a planned experiment. As statistician George Box remarks, to find out what happens to a system when you interfere with it you have to interfere with it, not just passively observe it.37

Evidence and Argument

The argument is the link connecting data and information with the conclusions of the analysis. As already noted, the structure of the argument typically will be a complex blend of factual statements, interpretations, opinions, and evaluations. Hence, whatever testing is done must rely on a variety of professional standards, corresponding to the different theories and methods employed; on the plausibility of the results and their robustness with respect to variation in the underlying assumptions and specifications; and on the criteria of adequacy of the client, or the rules of argument prevailing in the relevant forum of debate.

The nature of the evidence plays a crucial role here, since a wrong assessment of its strength and fit before it is included in the argument can lead to pitfalls in the drawing of conclusions. Even a style of presentation that is inappropriate for the audi-

36. K. Holden, "The Effect of Revision of Data on Two Econometric

ence to which the argument is directed can destroy the effectiveness of information as evidence.

Among the most widespread pitfalls related to evidence and arguments, three deserve special notice. The first originates in the contemporary fashion of using mathematical formalizations on every possible occasion. Two experienced analysts observe that "the analyst often dresses up his results and attempts, either consciously or unconsciously, to hide fairly elementary notions in extreme mathematical and technical language. Though it is probably not possible to condense the most esoteric results of modern mathematics and physics into the language of the newspapers, this is just not true of any applied operations analyses that we have seen."38

It should be added that an overly formalized style of presentation not only obscures the real issues and impedes assessment of the plausibility of the conclusions; it also induces a tendency to accept statistical information or the results of mathematical calculations as facts rather than evidence.

The second group of pitfalls is encountered when existing information is taken over for use in an analytic argument. All kinds of distortions occur when data gathered by one organization for broadly defined purposes are used by others to support specific conclusions. Whether such material is of sufficient strength and fit for its function in the argument depends on the mode of its original production; this is often difficult for the analyst to assess and usually impossible for him to change.

Finally, questions concerning the acceptable degree of approximation of numerical results or the acceptable level of precision of a set of data acquire their full meaning, for policy analysts at any rate, in connection with the use of evidence. Two other pitfalls should be mentioned in this context: the belief that there is an absolute standard of adequacy, and the rejection of items of information or opinions for which consensus among the experts is lacking. The belief in absolute standards overlooks the fact that even the physical sciences simultaneously use several degrees of acceptable levels of precision for their data. For example, some physical constants are known with an accuracy of $10^{-14}$, while the age of the earth can only be estimated with an error of billions of years. Because of the diversity of the data used in a typical analytic study, the acceptable margins of error may have to be even larger than those the economist or the sociologist must realistically accept. This does not mean, of course, that the analyst should not have high standards of quality for his evidence; the pitfall consists in setting the standards so high that they become self-defeating.

Conclusions

The conclusion of a policy study may be a prediction, a recommendation, an evaluation of ongoing programs, a new proposal, or a different perspective on an old problem. Whatever its nature, a conclusion always depends on a number of assumptions and methodological choices. A different conceptualization of the problem, other tools and models, or a few different judgments made at crucial points of the argument could lead to quite different conclusions.

Thus, the contact with the external world of real people and their problems is always indirect and elusive. This is true of any kind of intellectual inquiry, including natural science. But in science the pitfalls encountered when a theory makes contact with reality can be detected, before too much harm is done, by various means—including controlled experiments and working models—that reduce the abruptness of the impact. In policy analysis such tests are seldom, if ever, available. How, then, can one control the validity of a conclusion, make sure that it is not fallacious?

To repeat, policy problems are not textbook quizzes; they carry no guarantee that there always exist correct solutions against which analytic conclusions may be checked. Unlike the analyses of military operations conducted during World War II, and some small-scale industrial and administrative applications, it is extremely difficult to evaluate the usefulness of large-scale policy studies in terms of actual results produced. This is due to a number of reasons: first, the long time lag between the

38. Kahn and Mann, Ten Common Pitfalls, 47.
adoption of a policy recommendation and its actual implementation; second, the difficulty of sorting out the effects of a particular decision from among a multitude of confounding factors; third, and most important, the fact that the political and institutional context in which policy studies are done, has changed considerably over the years. In the early days of operations research and systems analysis the relationship between decision maker and adviser, between producer and user of analysis, was clearer and more direct than it is today. Now it is quite common for policy research to be sponsored by one organization, carried out by another, utilized by a third organization, and perhaps evaluated by yet another agency (which, in turn, may entrust the evaluation to an independent research group). Clearly, the criteria of effectiveness of the sponsors are not the same as those of the users, or of the controllers. Thus, analysts must satisfy a number of different, sometimes conflicting, expectations. The best they can do is to achieve some acceptable level of adequacy along each dimension; they must "satisfice" rather than maximize any one particular criterion.

CRITERIA OF ADEQUACY

We have shown that the work of the analyst is guided and controlled by many informal judgments concerning the different aspects of the problem under investigation. The precepts of a craft can never be fully articulated, but despite the absence of explicit rules a connoisseur has little difficulty in distinguishing good from poor craftsmanship. This is because the true connoisseur has an intimate knowledge of what might be called the microstructure of the artist's style.

Similarly, in order to appreciate the craft aspects of analysis and to evaluate competently the quality of the finished product one must learn to examine the microstructure of arguments. This was the purpose of breaking down the analyst's task into its component elements. Such a detailed examination would have only academic interest if it were possible to assess the quality of a policy analysis simply by comparing its conclusion with policy outcomes. A single synthetic criterion would be sufficient in this case: the analysis is good if the policy succeeds, and bad otherwise.

But, as I have said, policy problems carry no guarantee that there exist correct solutions against which the analyst's conclusions could be checked. Policies usually fail in some respects and succeed in others, and the relationship between what the analyst does, or says, and the final outcome is always indirect and uncertain. One recurrent theme in the evaluation literature is the call to develop methods that emphasize outcome rather than process. Outcome-oriented evaluation has a strong intuitive appeal, but in the case of professional work the outcome/process dichotomy is not very useful. A well-known medical expert has stated that "much, perhaps most, of what a physician does must be categorized as process, and process not even calculated to affect outcome." To a greater or lesser extent the same is true of all professionals, including policy professionals.

Hence process-oriented criteria of adequacy are necessary, although not sufficient, to assess the quality of analysis. A policy analysis is adequate if it meets the particular tests that are appropriate to the nature, context and characteristic pitfalls of the problem. Different criteria apply to the separate elements of the analysis, as exemplified in the following table:

<table>
<thead>
<tr>
<th>ANALYTIC COMPONENT</th>
<th>CRITERIA OF ADEQUACY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>Reliability, reproducibility, credibility (for exogenous data)</td>
</tr>
<tr>
<td>Information</td>
<td>Relevance, sufficiency, goodness of fit, robustness</td>
</tr>
<tr>
<td>Evidence</td>
<td>Reliability, admissibility, strength</td>
</tr>
<tr>
<td>Argument</td>
<td>Cogency, persuasiveness, clarity</td>
</tr>
<tr>
<td>Conclusion</td>
<td>Plausibility, feasibility, acceptability</td>
</tr>
</tbody>
</table>

At the beginning of this chapter I asked how it is possible to define standards of quality, and thus avoid falling into methodological anarchism, once the claim to certainty of conclusions is abandoned. Process-oriented tests of adequacy are the answer, or at least part of the answer. Policy analysis can be neither performed competently nor used properly without an appreciation of its craft aspects.

Naturally, avoidance of pitfalls and other tests of adequacy only guarantee minimal standards of quality. They do not, and cannot, imply originality, depth, or any other of the intellectual qualities that distinguish the brilliant from the merely competent study. Nor do they guarantee that the analysis will be useful to the people who pay for it, or that it will have an impact on the public debate. The question of the utilization of knowledge in the policy process is considerably more complex and will be taken up in a subsequent chapter.

To try to do something that is inherently impossible is, to borrow from Oakeshott, always a corrupting exercise. However, the tendency to equate the desirable with the feasible is always strong, especially in politics. As another philosopher has remarked, the existence of social tasks that appear both desirable and feasible but are in fact impracticable has set the stage for a wide range of conflicts in modern history. All the battles of social reform were fought partly on these grounds, with conservatives overstating and progressives underestimating the limit of the possible in public policy.¹

Helping policymakers and public opinion avoid both harsh overstatement and reckless underestimation of those limits is one of the most useful contributions analysts can make to public deliberation. A competent feasibility analysis attempts to identify all actual or potential constraints, separate them from fictitious obstacles, evaluate their significance for different implementation strategies, and estimate the costs and benefits of relaxing those constraints that are not absolutely fixed.