The Critical Appraisal of Scientific Inquiries with Policy Implications

William C. Clark and Giandomenico Majone

Since the Second World War, scientific information and analysis have been sought increasingly as aids to the resolution of practical policy problems. This growing demand for usable science has encouraged a rapid increase in the supply of would-be scientific policy guidance and advice. Much has been accomplished; but a growing tide of critical commentary indicates that all is not right.[1] As Lindblom and Cohen have characterized the situation, analysts are dissatisfied "because they are not listened to," while policymakers are dissatisfied "because they do not hear much they want to listen to."[2]

Many factors influence these laments, the most obvious being ignorance and conflict. For in practice, scientific inquiry cannot discover most of the things that policymakers would like to know. Much of what it does discover remains uncertain or incomplete. How, then, is "scientific" knowledge any more reliable a guide to policy than other forms of knowledge, prejudice, or propaganda? Moreover, experts often disagree on what science knows and on what that knowledge means for policy. If the knowledge produced by science is not consensual, what special claim for hearing can it make in a world of multiple opinions and biases?

Among the many prescriptions offered for mitigation of such questions in the last 25 years, the repeated calls for strengthened mechanisms of peer review and critical evaluation stand out. Practitioners and philosophers agree on the general principle: No rational enterprise can succeed without an effective critical capacity. Stephen Toulmin has pointed out that, in the absence of accepted and effective critical standards, judgments are clouded by:

> factors such as conservatism or prejudice, lack of professional cohesion or breakdown in communication, political pressure or sheer inattention . . . considerations which are entirely irrelevant to the problem under debate. . . . As a result, the disciplinary merits of some new terminology, technique of representation, or method of explanation may for the time being be disregarded, despite the fact that they could make themselves evident—given 'daylight and fair play' . . .[3]

Mere opinion and propaganda accumulate. The true limits of knowledge are obscured. Unproductive conflict is inevitable and uncontrollable.

The difficulty with the mechanisms of peer review is not in the principle, however, but rather in the practice, especially when the science in question has policy implications. The unproductive and uncontrolled conflicts symptomatic of an underdeveloped critical capacity remain a central feature of contemporary efforts to evaluate scientific advice. These efforts resemble less a rational discourse than a turkey-shoot (although it must be admitted that the producers of would-be scientific advice have provided plenty of bona fide turkeys at which to shoot). Study after study gleefully demonstrates that scientific inquiry in policy contexts is shot through with "fatal" methodological flaws, "hidden" biases, erroneous data, or trivial intent.

Much of this criticism is well deserved. At the same time, however, it is the kind of criticism that, in other contexts, would eventually trap itself

Dr. Clark is a research scholar at the International Institute for Applied Systems Analysis, A–2361 Laxenburg, Austria; Dr. Majone is a visiting professor at the John F. Kennedy School of Government, Harvard University, Cambridge, MA 02138.

Science, Technology, & Human Values, Volume 10, Issue 3, pp. 6–19 (Summer 1985) CCC 0162-2439/85/030006-14$04.00

in such absurd judgments as "Shakespeare can't write" or "Picasso can't paint." In other words, no comprehensive appreciation of what constitutes "good work" underlies the critical tradition now being applied to science in policy contexts. Instead, parochial sniping picks off one study for not being sufficiently rigorous to be real science, another for not being sufficiently open to be socially legitimate. Missing is any indication that good scientific inquiry in policy contexts might have more appropriate objectives than the emulation of either pure science or pure democracy.

More effective evaluation of the implications of ignorance and conflict for science in policy contexts will surely require that the critical standards currently arrogated from both classical physics and Jeffersonian democracy be replaced by standards more consciously and intelligently tailored to the task at hand. In light of this need, surprisingly little attention has been devoted to the development of an appropriate theory of criticism for scientific inquiry conducted in policy contexts. To be sure, a number of studies have adopted critical attitudes. Many of these might be called "rational criticism," focused on the development and use of predictive models and other computational techniques in policy contexts.[4] Another tradition of "practical criticism" focuses on the evaluation of policy efficacy per se.[5] Finally, the field of "ethical criticism" deals with the proper role of scientific knowledge in society.[6]

There is nothing wrong with these individual critical perspectives. Much can be learned and many difficulties could be avoided by viewing today's problems of applied scientific inquiry from their vantage points. Nonetheless, as Northrop Frye has remarked in the context of literary criticism, there seems to be no reason why the larger edifice to which these individual perspectives are contributing should remain forever invisible to them, as the coral atoll to the polyp.

Comprehensive criticism should also be possible. It would recognize the validity of existing critical perspectives but would also seek—by making those perspectives more aware of one another—to construct a critical vision that is more than the sum of its rational, practical, and ethical parts. Our goal in this paper is to make a start in the formulation of such a comprehensive critical perspective.

We begin with a question that is central to discussion of peer review systems: "Criticism by whom?" We emphasize the importance of accounting for the different *roles* in the process of critical evaluation. The next section turns to the question of critical modes—i.e., "Criticism of what?" It examines the implications of choices to evaluate inquiry in terms of the inputs, the outputs, or the process by which inquiry is conducted. We then use the concepts of role and mode to help provide a classification of the critical *criteria* that are used in actual practice. This approach leads to a discussion of comprehensive metacriteria relating to the *adequacy, effectiveness, value,* and *legitimacy* of scientific inquiry performed in policy contexts. We close with some tentative suggestions on the practical implications of a comprehensive critical perspective for the construction of usable scientific knowledge.

Critical Roles

Scientific inquiry in policy contexts can serve a wide variety of uses and users. The same inquiry may be used by other scientists as a foundation for subsequent study, by policymakers to guide their choice of actions, by program managers making funding decisions, or by reporters trying to inform the public. What is "good" or "useful" inquiry for some of these users may not be so for others. Before we can speak sensibly about the design of criteria for the critical appraisal of scientific inquiries conducted in policy contexts, we must therefore face another question: Which interests or roles can be regarded as appropriate sources of those criteria? Who are, in fact, the "peers" who will be doing the reviews? It is clear, for example, that an academic statistician, a company expert, and a Congressional staffperson will use different criteria to evaluate a study on the correlation of cancer deaths and particular occupations. Their conflicting judgments of the study will not be resolved by reference to less uncertain or more clearly presented data. Required instead is a mutual comprehension of the different critical perspectives being employed.

That different critical appraisals are arrived at by people in different roles is not a bad thing as such. It may simply reflect different needs and concerns of different segments of society, or different degrees of freedom in making certain key methodological choices. So long as the judgments leveled from the perspective of one particular role are not presented or misinterpreted as judgments relevant to or speaking for all possible roles, we have a healthy state of pluralistic criticism. Difficulties begin to arise when this neat partitioning of roles and the criticism voiced from them begin to break down. Unfortunately, such breakdowns seem more the rule than the exception in actual practice.

Perhaps the most common problem occurs when an inquiry designed for use in one role receives its only critical review from the perspective of a different role with different critical standards. Because the standards are mismatched, the results of the inquiry are almost inevitably found wanting. Such difficulties arise, for example, when academic ecologists are convened as the sole reviewers of the environmental impact assessment mandated for Federal projects. From the perspective of a project manager or administrative law judge, the relevant critical criteria might be the timeliness of the assessment, the likelihood that potentially serious impacts have been noted, and, perhaps, whether practical development alternatives are suggested. From the ecologists' point of view, however, such criteria are (at best) vaguely comprehended and given only secondary consideration. Their criticism would likely be based on such criteria as the adequacy of sampling design, the use of appropriate theory, and the accurate characterization of uncertainties. As a result, the ecologists may accept or reject attempted impact assessments for reasons largely irrelevant to the people who will eventually have to resolve the practical problems of environmental management.

A third difficulty appears when well-intentioned efforts to honor the perceived critical standards of one or another group result in scientific inquiry delivering more than it knows. This point was clearly illustrated in the hearings of the Joint Economic Committee of the U.S. Congress, held in February 1975 to provide scientific advice on the question "What is the size of America's remaining oil and gas resources?"[7] Senator Hubert Humphrey called the hearings in response to a National Research Council (NRC) report that had reviewed previous estimates by industry and government experts and then produced its own scientific findings. The NRC showed that serious scientific studies of the question produced estimates spanning an order of magnitude, and that government (U.S. Geological Survey) estimates tended to be two to three times the size of most industry estimates. The NRC itself concluded that the most reasonable estimate was less than half the most recent Geological Survey figure, and only slightly larger than those proposed by industry experts.

Senator Humphrey was not amused. Questioning the NRC and other experts called before his committee, he lamented:

> You cannot imagine, gentlemen, what hits us when these come out. . . . The mail is incredible. . . . One (caller) said to me, 'Didn't you read that report that came out from that group of scientists; aren't you frightened?' . . . . Now, help me. Where do you come down in this wide range of estimates? Do you feel it is the upper or the lower end or where is it?[8]

The Senator's appeal for the scientific facts was met with the following explanation by an NRC committee member of how the committee has reached its own estimates: ". . . estimates of future supplies of oil and gas are so dependent upon unknown scientific factors and unknown environmental and political factors as to be almost unknowable." These "almost unknowable" estimates were nonetheless published to three significant figures by the NRC with no uncertainty ranges. How were the particular NRC values arrived at? According to another NRC member, "from our point of view, we though it advisable . . . to accept more conservative estimates, thinking that most of the Geological Survey estimates are relatively high, and most of the oil company estimates relatively conservative."[9] As Wildavsky and Tenenbaum ask in their review of the case, "This is science?"[10]

Congress wanted "a number" and the National Research Council gave them "a number," even though committee members acknowledged in their testimony that it was little more than guesswork—i.e., nothing like the consensually certified knowledge that its trappings and origins implied. Similar examples, of which perhaps the most notorious would involve the willingness of scientists to deliver cost-benefit assessments of long-term and large-scale environmental changes, could readily be cited.

Is there a cure for such common tendencies to confound critical roles? Probably not. At a minimum, however, efforts to build a critical capacity for judging scientific inquiry in policy contexts should explicitly recognize that multiple roles ("peer" groups) exist, each with a legitimate claim to set critical criteria. Further, such efforts should appreciate the complex pulls and pushes that the resulting diversity of critical criteria will exert on the inquiry itself.

For the appraisal of science conducted in policy contexts, the minimum set of roles to consider would probably include the individual scientists performing the inquiry, their disciplinary peer groups, the sponsor or manager of the research program, the client or decisionmaking group for whose use the results of the inquiry are intended, and, in most cases, some version of the interest groups that could be expected to have a stake in decisions being contemplated.

In principle, it might be possible to envision a grand scheme that would combine all the role perspectives into one common critical standard, thus providing a weighted evaluation of any scientific inquiry in its appropiate policy context. Some decision analyst has probably already proposed such a scheme, but we are grateful not to have seen it. Technical difficulties aside, we suspect that a common standard, like any other aggregational procedure, would solve none of the important problems and would create new ones of its own. Our own suspicion is that efforts to develop better critical skills for science with policy implications should aim not for a unique evaluation, but rather for an enhanced understanding of different evaluative criteria on the part of all role players. We most need, in other words, a more sophisticated and sympathetic understanding of the multiple perspectives involved.

Critical Modes

Role considerations are not the only factors that need to be distinguished in making sense of the criteria by which scientific inquiries in policy contexts can be critically appraised. Analysts have also found it useful to distinguish three general *modes* of critical appraisal. In the *output* mode, appraisal focuses on the products of inquiry: scientific facts, problem solutions, conclusions, and the like. In the *input* mode, the emphasis is on the data, methods, and people engaged in the inquiry. Finally, in the *process* mode, attention shifts to institutional structures and procedures, provisions for quality control, and questions of standing that govern participation in and conduct of the inquiry. We next consider each of these modes individually, all the while keeping in mind that they are often mingled in practical efforts to appraise scientific research and advice.

Appraisal by output or results is, in a commonsense view of criticism, the obvious way to assess the "goodness" of any rational enterprise. Goals are defined, results are produced, and the two are compared. Can the hypothesis be rejected? Is the problem solved? Answers to such questions are often all that is needed for critical appraisal. But the intuitive appeal of this mode of criticism is often misleading. It is also important to emphasize the circumstances under which appraisal by outputs is likely to be impossible, ineffective, or at least more difficult than appraisal by other modes.

As one example, the output mode of appraisal cannot function when goals are ambiguous or undefined. It works only slightly better when goals are multiple or contested. Yet, one of the most important realistic goals that can be pursued by scientific inquiry in policy context is to provide general "enlightenment."[11]

Appraisal by results also fails when the outcome or results of a scientific inquiry with policy implications cannot be known or measured with sufficient certainty within a meaningful time frame. This situation arises commonly in practice, as suggested by Haefele's "hypotheticality" dilemma for studies of nuclear reactor risks,[12] Weinberg's discussion of the "transscientific" status of studies on the biological effects of ionizing radiation,[13] and the problems of validation or detection facing most large simulation modeling exercises.[14]

A strictly output-directed appraisal leaves no room for dealing with such difficult cases. Nor does it let us escape an equally awkward utilitarian posture with respect to means. Exclusive reliance on outcome appraisal implies that the means are neutral, with only instrumental values. But means are often valued as such. And what if the inputs to the analysis—for example, the experimental protocols for treatment of test subjects—are unacceptable? An approach to the appraisal of practical scientific inquiry requires modes of criticism that can free it from this utilitarian trap.

Having noted these shortcomings of the output mode of criticism, it is easier to see the need for, and uses of, alternative and complementary modes. Appraisal by input is concerned with the actual activity of inquiry. Who is doing the analysis and what is his or her track record? How good are the data and how competently are they used? What is the overall maturity of the intellectual disciplines involved? How much time and support is available? Appraisal by inputs clearly has its own dangers. As amply illustrated by the Congressional budget process in the United States, when inputs are but loosely attached to outputs, great battles over how much goes into a program can be won and lost without much affecting the problem-solving in the outside world. Nonetheless, because of the frequent difficulty of measuring outputs, and the expense in time and other resources of doing process evaluation (see below), appraisal by input is often all that the critic has to work with. Moreover, to the extent that one purpose of critical appraisal is to guide an ongoing inquiry (rather than to grade a completed one), it is by attention to inputs that control can best be exercised.

Finally, appraisal by process represents an additional alternative to the commonsense mode of output appraisal. One of the most important functions performed by actual scientific inquiry in policy contexts is the provision of procedures for the resolution of partisan arguments. Such a function can be appraised only in terms of procedural considerations. "Due process" is appealed to in a wide variety of practical problemsolving circumstances, including those encountered in legal and administrative thinking, in scientific practice, and in a variety of learned professions like medicine. In all these cases, the appeal to process is made in order to enable critical appraisal of findings, decisions, or actions in the absence of accepted criteria of truth, justice, or practical effectiveness. Moreover, a sharp distinction of the outcome and process modes of evaluation protects the practitioner against the risks inherent in professional practice. A useful critical theory must be able to evaluate the extent to which such procedural considerations have actually been met. Of the three critical modes, evaluation by process is the most subtle and informative—it provides information that input or output measures are almost sure to miss. But it is also the most costly, because it requires intimate knowledge of the process and direct, extensive observation. It cannot rely on simple statistics and aggregate data.

Critical Criteria

The foregoing discussion of critical roles and critical modes provides a general framework from which it is possible to classify the diverse critical criteria used in actual cases of scientific inquiry in policy contexts.

Our approach to this classification has been empirical. Rather than beginning with one narrow perspective of what appropriate criteria "should" be, we have surveyed a variety of historical efforts to evaluate "systems" studies of scientific investigation conducted in policy contexts. The kinds of criteria adopted in those studies are summarized in Table 1. We emphasize that this is merely a preliminary sketch. A more comprehensive review, drawing more extensively from casework in decision and policy analysis per se, would doubtless be more informative. In particular, the focus of our survey on environmental and energy problems probably overemphasizes the importance of criteria important to "public interest" roles, and underemphasizes the more sophisticated scientific criteria that would be expected to emerge from, say, a review of pharmaceutical safety studies. Nonetheless, the limited material presented in the table is a beginning.

The entries in Table 1 suggest the daunting variety of critical criteria being employed in current practice. The very diversity of standards in the table helps us to understand some of the conflicts that arise over the appraisal of science in policy contexts. For example, the charade of the Congressional hearings on oil and gas reserves can be seen as an instance in which the National Research Council committee, acting in a peer group role, adopted a largely process-oriented mode of critical appraisal in guiding its own work: "We thought it advisable . . . to accept more conservative estimates. . . ." Individual scientists on the committee, on the other hand, implicitly adopted an input-oriented mode when acting in their personal roles as researchers: "Estimates of future supplies . . . are so dependent on unknown scientific factors . . . as to be almost unknowable." Meanwhile, Senator Humphrey, berating the study in terms of the output mode of criticism most important to him in his policymaking role, asked: "Where do you come down in this wide range of estimates? Do you feel it is the upper or the lower end or where is it?" From the table, it is clear that there existed virtually no overlap in the critical criteria underlying the three different evaluations of the oil and gas studies. Whatever more fundamental disagreements may have existed regarding the worth of the studies, most of the conflict in the hearing room can be traced to the different critical standards being employed by different critics. Rectifying the shortcomings perceived by the scientists would not have improved the reception of the studies by Senator Humphrey. And efforts to meet the Senator's criteria by pro-

Table 1. Critical criteria.

Critical Role	Critical Mode: Input	Critical Mode: Output	Critical Mode: Process
Scientist	Resource and time constraints; available theory; institutional support; assumptions; quality of available data; state of the art.	Validation; sensitivity analyses; technical sophistication; degree of acceptance of conclusions; impact on policy debate; imitation; professional recognition.	Choice of methodology (e.g., estimation procedures); communication; implementation; promotion; degree of formalization of analytic activities within the organization.
Peer Group	Quality of data; model and/or theory used; adequacy of tools; problem formulation. Input variables well chosen? Measure of success specified in advance?	Purpose of the study. Are conclusions supported by evidence? Does model offend common sense? Robustness of conclusions; adequate coverage of issues.	Standards of scientific and professional practice; documentation; review of validation techniques; style; interdisciplinarity.
Program Manager or Sponsor	Cost; institutional support within user organization; quality of analytic team; type of financing (e.g., grant vs. contract).	Rate of use; type of use (general education, program evaluation, decisionmaking, etc.); contribution to methodology and state of the art; prestige. Can results be generalized, applied elsewhere?	Dissemination; collaboration with users. Has study been reviewed?
Policymaker	Quality of analysts; cost of study; technical tools used (hardware and software). Does problem formulation make sense?	Is output familiar and intelligible? Did study generate new ideas? Are policy indications conclusive? Are they consonant with accepted ethical standards?	Ease of use; documentation. Are analysts helping with implementation? Did they interact with agency personnel? With interest groups?
Public Interest Groups	Competence and intellectual integrity of analysts. Are value systems compatible? Problem formulation acceptable? Normative implications of technical choices (e.g., choices of data).	Nature of conclusions; equity. Is analysis used as rationalization or to postpone decision? All viewpoints taken into consideration? Value issues.	Participation; communication of data and other information; adherence to strict rules of procedure.

viding more precise numbers would almost certainly have resulted in rejection of the study by the scientists. Tinkering with methods for the better portrayal of uncertainty or the inducement of greater consensus would not have solved anyone's problems.

The approach of understanding differences in critical perspectives suggested in Table 1 can be used in other diagnostic situations and might help to improve the quality of debate over uses of science in policy contexts. Given the multiple critical perspectives emerging from the matrix of roles and modes, however, the question remains whether any more general statements can be made about the nature of critical criteria appropriate for use in such situations. This would be the case if the various criteria and subcriteria suggested in Table 1 could be usefully grouped into a small number of metacriteria which, at a general level, cut across considerations of role and mode. In the following sections, we argue that such metacriteria do exist, and sketch them under the headings of adequacy, value, effectiveness, and legitimacy.

Criteria of Adequacy

Scientific inquiry in policy contexts is plagued by problems of "ineffectiveness." It is unable to accumulate the certified "facts" that provide the "grounds for peaceful discourse."[15] The most obvious (although not the only) cause of this ineffectiveness is systematic weakness in the materials of data and methods with which applied scientific inquiry must work, and in the mortar of inference and argument with which those materials must be combined. It is to shore up such weaknesses that we look to critical criteria of adequacy. Jerome Ravetz, in his extensive study of the subject, argues that the function of criteria of adequacy is thus to make "facts" possible.

Criteria of adequacy include most of the commonsense notions of evaluation with which scientists deal in their everyday activities. Thus, it is second nature to most scientists to question origins and reliability of data, to test apparent correlations for the effects of chance, and to demand that test procedures be replicable. The existence of communally accepted facts is fundamentally dependent on the belief that the original investigator and the peer groups reviewing his or her work have considered such criteria of adequacy.

Looked at more deeply, however, two distinct uses of adequacy criteria emerge. The first, notes Ravetz, is to channel critical disputes and debates to well-defined categories where focused discussion and rational resolution can be carried out. In the absence of such channeling, "controversies on results range indiscriminately and inconclusively from criticism of raw data to abstract methodology."[16] Most scientists engaged in policy-relevant work have experienced the frustration of such uncontrolled debate where rhetorical style counts for more than technical accuracy. Worse, many have found that their efforts to establish "ground rules" (i.e., criteria of adequacy) before the debate commences are viewed with suspicions or downright hostility. And the suspicions come not only from their opponents but also from the public at large, from administrative law judges, and from members of Congress.

Once again, these are not the sorts of difficulties that can be remedied by a new computational method or hearing procedure. Rather, they reflect a profound need for better education throughout society concerning the limited competence of scientific inquiry in policy contexts. Moreover, they demonstrate the requirements for appropriate and critical standards if the inevitable mistakes and blunders of scientific inquiry are to be identified and eliminated.

A second use of adequacy criteria is to prevent scientific blunders from occurring in the first place. In this context, however, it is important not to equate criteria of adequacy with rules for good practice. The vulgar version of the "scientific method" taught in schools (and formalized in such "scientific" procedures as environmental impact check lists) presents just such rules: "Do it this way, test the results, and the resulting knowledge will be scientific." Nothing, of course, could be further from reality. Scientific inquiry involves creativity plus essential elements of craft skill. There is no "standard operating procedure" that will guarantee good science, any more than there is one that will guarantee good painting. In both cases, however, there are pitfalls that most good practitioners learn to avoid.

In performing basic statistical operations, for example, it is a sign of naïveté or incompetence to attempt some of the more conventional significance tests without testing for normal distribution of the data. The same is true even in descriptive statistics when, say, the arithmetic mean is used instead of more appropriate parameters like a geometric mean or a median. An important purpose of adequacy criteria is to recognize such pitfalls and to "sign post" them in a way that makes avoidance possible by any reasonably attentive practitioner. If such "sign posting" is not done, much work will be vitiated by pitfalls, while the rest will be of doubtful utility simply because of the suspicion that undiscovered pitfalls may have compromised the analysis. Once again, the existence of suitable criteria of adequacy is a necessary condition for the accumulation of credible facts.

This is not the place to explore in any detail the specific pitfalls associated with scientific inquiry in policy contexts. The beginnings of such an exploration are given in Quade's *Analysis for Public Decisions*, and Majone and Quade's *Pitfalls of Analysis*.[17] The systematic development of specific adequacy criteria to avoid these and other known pitfalls of present practice would make a logical and useful follow-up study to the groundplan developed here.

Criteria of Value

To serve its intended function, criticism must provide more than post hoc evaluations of the adequacy of completed scientific inquiry. In addition, it needs to facilitate prospective judgments on which scientific inquiry to undertake. Criticism applied to such questions of scientific choice involves criteria of value. The word "value" is used here in the sense of "worth" or "potential," as in the "value" of a baseball player for his team's pennant hopes.

For the case of science in policy contexts, questions of choice become enormously complex: scientists choose what studies to perform, institutions choose what work will be on their agenda, program managers choose what research to fund, and policymakers choose what problems to tackle (and in what order). Without appropriate criteria of value to guide and evaluate such choices, ineffective and unproductive undertakings almost inevitably result. For scientists, even for those manifestly concerned that their work be "relevant," there is a great tendency to take on what John Passmore has called a "charmed circle" of problems, rather than to draw from the full range of challenges posed by the real world. For politicians, there is a strong temptation to go for easily conceptualized, plausibly "urgent" headline-grabbers. For research institutions, there is inexorable pressure to undertake whatever work will increase the budget. So-called peer elite groups seem destined to put on their lists of research priorities a little work for each component discipline. Cutting across all these roles in the science-for-policy play, there is a strong inclination to demand, promise, and undertake tasks far beyond the present capacities of scientific inquiry.

In contrast to this litany of choice-distorting pressures, the desirable situation would be one in which the choice of a scientific inquiry with policy relevance had something to do with the "importance" of the problem addressed. In other words, low marks would be given for patently trivial proposals, whether the triviality involved matters of scientific fact or social concern. In addition, however, desirable choices would reflect some notion of feasibility. For if, as Medawar has argued, good science is the "art of the soluble," then good science-for-policy should surely have something of the same character, however urgent the practical problem of interest.

Part of the challenge of building more effective critical skills for science in policy contexts is to make more explicit such commonsense notions of value. Problems of scientific choice and their related criteria of value have been little studied by philosophers of science, preoccupied as they have been with the finished products of scientific inquiry. Instead, most of the discussion has come from people engaged in the setting of policy for science and in the allocation of research funds. Much of the best work has appeared in *Minerva* over the last 20 years and is summarized in the volume on *Criteria for Scientific Development*, edited by Edward Shils.[19] Two classic papers from that volume by Weinberg and a long commentary on them by Ravetz form the basis of the summary account of scientific value criteria given here.[20]

Alvin Weinberg discusses two basic components of value criteria, which he defines as "internal" and "external." Jerry Ravetz introduces an important third component, the "personal." Internal criteria of value are addressed to the question: "How well is the scientific inquiry under consideration likely to be carried out?" Answers to such questions must clearly come from other scientists within the discipline. Relevant considerations will be the quality of the people in the field in general, and the reputation of the particular researchers likely to be involved in the project. In addition, attention will be paid to the maturity of the field itself in relation to the tasks set before it. How effective has it been in the past? Are its goals and problem areas well defined? Has it developed sufficiently strong criteria of adequacy to ensure that debates will be productive and facts established? Or, in Toulmin's terminology, is the field "well disciplined?"[21] Internal criteria of value, in other words, reflect the extent to which the proposed inquiry has come to terms with "the art of the soluble."

There is nothing particularly remarkable about the internal criteria of value described above. This is how most scientists would see their everyday activities of evaluating grant proposals, sitting on committees to establish research priorities, and so on. Weinberg's insight was to recognize that such traditional internal criteria of value are unable to help evaluate science's position within, or its obligations to, a larger social context. To provide that larger critical context, he argued, required additional *external* criteria of value.

The key feature of external criteria of value is that they are generated and applied not by the experts in the science being evaluated but rather by people outside the evaluated field of inquiry who will use its results. Such users may be other scientists in related disciplines, problem-solvers looking for technical help, or a society-at-large looking for general enlightenment or ritualistic reassurance. In all cases, the central question is the extent to which the goals of the inquiry, if achieved, would contribute to the solution of problems *outside* the field of science doing the research.

Finally, Ravetz has called attention to a class of *personal* criteria of value ignored by Weinberg. Ravetz's argument is that, especially where science bears on practical problems, the personal concerns of an individual researcher may be a significant factor influencing scientific choice. This seems to us an enormously important point. Studies in the sociology of science, as well as everyday experience, show how important the choice of research topics is in determining the current state of scientific knowledge. In fact, as Ravetz has pointed out in a more recent paper, the largest single determinant of ignorance and uncertainty in many fields of practical significance may be the choice by individual scientists *not* to perform (or, of course, of individual program managers *not* to fund) research in those areas.[22] Personal criteria of value therefore can be extremely important in determining which problems are illuminated by vigorous scientific inquiry, and which are not. Correspondingly, a critical theory seeking to evaluate scientific choice must have a means of conceptualizing the play between such personal valuations, and the more objective valuations applied by other people internal and external to the field of inquiry.

The subjective element present in any critical judgment is especially important in evaluating the individual components of criteria of value. Combining personal criteria with those imposed from internal and external sources is even more problematical. Precisely for these reasons, however, Ravetz has stressed that "an attempt to eliminate the elements of craft experience and personal wisdom in these judgments and substitute for them a bureaucratic routine would soon produce gross errors of planning. . . ."[23] The people most intimately involved in doing scientific research, or in using the results of that research in problem-solving and policymaking contexts, must therefore be central to the formulation and assessment of criteria of value. The institutional difficulties of bringing these role players together in a critical undertaking are substantial. But the importance of getting on with the task is equally great.

Criteria of Effectiveness

Does scientific inquiry actually help to resolve practical problems? We have already suggested that systematic weaknesses in the objectives and materials of inquiry often fail to establish consensual "facts," thus leading to conflicts that render the results "ineffective." To address this difficulty more constructively will require the development of appropriate critical criteria of effectiveness.

Efforts to develop such criteria must contend at the outset with what Carol Weiss, commenting on problems of policy evaluation, has called the problem of "little effect."[24] With depressing regularity, evaluations of policies—whether they are designed to improve students' learning, the military's readiness, or the environment's health—produce verdicts that the world "out there" has remained pretty much the same. If policies themselves seem to be so ineffective in changing the world, what kind of effectiveness can we reasonably expect from the scientific inquiries that seek to influence those policies?

Part of the "little effect" problem is doubtless real. Most policy, like most scientific inquiry conducted to assist policy development, really *is* ineffective. Incompetence, like ignorance, is the rule rather than the exception. In addition, however, the finding of "little effect" is an inevitable outcome of inappropriate appraisal in the output mode that we discussed earlier. Ill-defined or controversial goals, lagged impacts, and the "noise" created by other policy initiatives can all induce a reading of "little effect" even when real change is afoot. Finally, however, the major cause of the "little effect" verdict seems likely to be a misconception of the policy process itself. Perhaps surprisingly, some of the most illuminating efforts to come to terms with this misconception of the policy process have their roots in the philosophy of science. By pursuing those efforts, we can establish a solid foundation for criteria of effectiveness in the intermediate ground of science conducted in policy contexts.

Thomas Kuhn, in *The Structure of Scientific Revolutions*, clearly identified the scientific equivalent of policy evaluation's "little effect." Kuhn noted that if, in line with the "falsificationist" thinking of positivist philosophers, "any and every failure of experimental results to fit were grounds for theory rejection, all theories ought to be rejected at all times."[26] But, in fact, he argued, most "normal" scientific experimentation has little effect on the theories it is nominally testing. In the words of Imre Lakatos, "nature may shout 'no' but human ingenuity . . . may always be able to shout louder. With sufficient resourcefulness and some luck any theory may be defended 'progressively' for some time, even if it is false."[27] Subsequent empirical studies in the history of science have made it clear that the "little effect" of most scientific experimentation is generally a good thing for science. Without it, most productive research programs would have been "refuted" to death in their infancy.

Observations of this sort led Lakatos and others to reformulate conventional interpretations of the development and critical appraisal of scientific inquiry. Lakatos proposed that the object of critical evaluation in science should be not the individual theory, but rather the "research program."[28] This he conceived of as historically evolving entity with a "hard core" of assumptions held temporarily beyond criticism, a "belt" of protective hypotheses that may be refuted by appropriate observation or experience, and a set of "heuristics" that guide the path of the program's development. The normal operation of science consists of a competition between alternative research programs, each with a somewhat different set of theories, goals, and predictive accomplishments.

Lakatos proposed to evaluate the effectiveness of a program in terms of the "progressiveness" of the problem shifts it induces. A problem shift is progressive if it has better powers of prediction or explanation than its predecessors. Thus, a research program is progressing so long as it keeps making novel predictions with some success; it is stagnating or degenerating when it can provide only post hoc explanations. Competing research programs are judged in terms of the relative progressiveness of their respective problem shifts. Although this is not the place to discuss the issue in any greater detail, the emergence and subsequent evolution of Lakatos's "research program" concept have brought both new vigor and new relevance to epistemological debates on criticism and the growth of scientific knowledge. Most relevant for our purposes, it has gone far towards finessing away the dilemma of "little effect" in critical evaluations of scientific inquiry.

Lakatos's concept of the evolving "research program" has been used to understand the comparably evolving nature of problemsolving efforts, and the effect of scientific inquiry upon them.[29] As Lakatos had redirected, the object of scientific criticism of science in practical contexts is not the individual decision or action but rather the larger policy or action program. In this view, "policy development is a sequence of partly overlapping action programs. The focus . . . [is on] objective features like policy content, evolving doctrines and problem situations, changing constraints, and interactions among different policies."[30] In terms derived from Lakatos's critical criteria, such an action program "may be said to be progressing as long as it succeeds in disposing of issues, i.e., in moving them from the stage of contention to a class of issues which the actors in the policy process judge to be in a state of satisfactory, if temporary, resolution."[31] The notion that effective policymakers must be able continuously (if temporarily) to remove selected issues from their agenda has long been central to evaluations of organizational or governmental strategies. As Lipset has put it, "moderation is facilitated by the system's capacity to resolve key dividing issues before new ones arise. If the issues . . . are allowed to accumulate, they reinforce each other . . . ," leading to unproductive conflict.[32]

The stage has therefore been well set for the development of criteria of effectiveness focused on the contribution of scientific inquiry to the control of policy *agenda*, rather than the policies themselves. Although the problems of "little effect" are not totally eliminated by this shift in perspective, any more than they were for science per se by Lakatos's contribution, indications are that the task of critically evaluating effectiveness should become markedly easier. Preliminary investigations, some of which were summarized in Table 1, show that this expectation is born out in actual practice.[33]

Criteria of Legitimacy

Two overriding questions asked regarding policymaking in open societies are its efficacy in solving practical problems and its responsiveness to popular control. As C. E. Lindblom remarks, however, these questions lead to

> a deep conflict [that] runs through common attitudes to policymaking. On the one hand, people want policy to be informed and well-analyzed. On the other hand, they want policymaking to be democratic. . . . In slightly different words, on the one hand they want policymaking to be more scientific; on the other, they want it to remain in the world of politics.[34]

The results of scientific inquiry performed in policy contexts are a potential source of political power. The question that arises, therefore, as it does regarding any source of power, is what constitutes its legitimate use. In appraising efforts to provide usable knowledge through scientific inquiry, we must consider criteria of legitimacy.

Weber introduced an elaborate typology of the modes and sources of legitimacy in a variety of social contexts.[35] But for our narrower purpose of coming to terms with the legitimacy of scientific inquiry in policy contexts, the division into "numinous" and "civil" sources of legitimacy may suffice. Briefly, numinous legitimacy is derived from superior authority held to be beyond questioning by those who endure the consequent exercise of power. Traditionally, the "god-king" doctrines of ancient and modern religions have been the most obvious examples. Additional cases are the inspirational authority of prophets and the other charismatic leaders. Civil legitimacy, in contrast, reflects a freely negotiated agreement or contract to follow certain rules or to consent to certain procedures. This is the common stuff of modern constitutional governments and economic systems. Clearly, it is through recourse to civil legitimacy that the "deep conflict" between expertise and popular control referred to at the beginning of this discussion must ultimately be addressed.

The social uses of science have always had something in common with the social uses of religion. And in the two decades following World War II, modern science took on almost religious-looking numinous legitimacy as an unquestioned source of authority on all manner of policy problems. But the unquestioning acceptance of science's legitimacy no longer holds. Comparing the present situation to that which developed in the wake of Sputnik, Harvey Brooks notes that "scientists today are listened to much more but believed much less than they were in those heady days. There is still a great respect for learning among politicians and policymakers, but there is also much greater skepticism and suspicion, and the image of objective, 'value-free' science and scholarship is severely tarnished."[36] Alvin Weinberg made the same point more directly in his call for scholars working in policy contexts to accept the "transscientific" nature of their work, and the consequent obligations for social negotiation.[37] What we see then, is that the postwar numinous legitimacy of science has been eroded, leaving in its wake a need for a socially negotiated civil legitimacy. Our society's great preoccupation in recent years with "public interest" and "critical" science, with hearing procedures and "independent" assessments, and with demands for "better" ethical standards of scientific practice reflect both the urgency and the difficulty of those negotiations.

Can studies of legitimacy in other contexts provide some guidance for the development of a socially acceptable and legitimated place for scientific inquiry in policy contexts? We believe so. Some of the recurrent critical quandaries encountered in efforts to make science both more useful and more controllable should yield to analysis from the perspectives that studies of political legitimacy have produced. For example, Guglielmo Ferrero, in his classic analysis of *The Principles of Power: The Great Political Crises of History*, argued that civil legitimacy rested on the two pillars of majority and minority, and on the arrangements made to control the treatment of the latter by the former.

The relationship between majority and minority is precisely the issue involved in the "fair play of ideas" that is the basis for progress in scientific inquiry. Difficulties in assuring such "fair play" abound, reflecting the frequent intolerance of established science for unconventional ideas. The phenomenon is particularly well analyzed from the point of view of science per se in Imre Lakatos's discussions of "monster-barring" strategies in pure science.[39] But the implications of how scientific inquiry is prepared to handle minority views when working in policy contexts has not yet been usefully explored at any deep conceptual level. An effort to integrate the "majority-minority" perspectives of the social scientists and the "monster-barring" perspectives of the science philosophers would therefore seem worth pursuing.

One of us has begun work in this direction through studies of the legitimizing use of scientific inquiry in the regulatory process.[40] And the general problem of providing institutional means for legitimate resolution of scientific conflicts with policy implications is currently being reviewed by Marc Roberts, Stephen Thomas, and their colleagues in an interesting program at the Harvard School of Public Health.[41] All this work emphasizes the close connection between political notions of access, standing, and agenda-setting on the one hand, and the traditional scientific values of tolerance and skepticism on the other. The empirical studies of Brian Campbell make it clear that science need not forfeit the legitimacy of its claims to special standing in technical policy debates, merely because it demonstrates uncertainty on particular factual and interpretative issues that arise.[42] Indeed, an increased willingness and ability to address more forthrightly the limits of science's competence and the extent of its ignorance are almost certainly prerequisites for enhancing a new and democratically negotiated civil legitimacy for scientific inquiry in policy contexts.

Conclusion

We have argued in this paper for a self-conscious, sustained effort to build a systematic critical perspective appropriate for the special needs of scientific inquiry in policy contexts. We have emphasized the need for a *comprehensive* perspective because the partial perspectives now in use inevitably slight the integrative and synthetic considerations so essential to useful inquiry on practical problems. Rational criticism, practical criticism, ethical criticism, and the like all have perfectly proper roles to play. But their full potential will never be realized until they also have a common framework that lets them stand back and see how each partial perspective can complement and reinforce the others.

We have also stressed the need for *appropriate* criticism because, in the vacuum created by its absence, scientific inquiry in policy contexts will continue to be criticized in terms of standards and criteria arrogated from pure science on the one hand and pure politics on the other. As we have argued, such inappropriate criteria can only distort the conduct of the scientific inquiries they seek to strengthen. We have suggested that much could be gained by focusing critical discussion through the four metacriteria of adequacy, effectiveness, value, and legitimacy. The need is to begin and sustain a wide-ranging dialogue on these criteria among the producers, users, and managers of applied scientific inquiry.

In addition to better understanding of the nature and ramifications of critical criteria, there is a need to develop a much wider and more effective array of institutional mechanisms for the exercise of critical judgment. Anonymous peer review will always have a role to play, but our analysis suggests that this role is of much more limited use in appraising science in policy contexts than it has been in treatments of science per se. Above all, there is the difficulty of assembling a single community of "peers" that will appear legitimate to both consumers and producers of would-be-useful scientific inquiry. Useful peer communities would have to develop expertise and sophistication in handling the whole range of critical criteria and considerations we have outlined here. Such experience is very rare today, and "peer" review of science in policy contexts may consequently be more a part of the problem than a part of the solution to producing more usable scientific knowledge.

Beyond peer review mechanisms, complementary critical fora are needed. The most difficult problem will probably be to engage appropriate people from outside the community of scientific researchers and program managers so that the difficult questions of value and legitimacy can be meaningfully reviewed. Institutional arrangements like the science court and the National Research Council's expert committees are demonstrably inadequate for the task at hand.

A priority goal should therefore be the exploration of new institutional mechanisms for critical appraisal and evaluation of scientific inquiry in policy contexts. At least in its early states, this exploration should be frankly experimental, adopting a variety of mechanisms in actual problem contexts and assessing the results.

We emphasize that this is not something that can be left to the policy analysis departments of universities or the comparable divisions of the Federal executive agencies. Rather, to be useful, the experiments will have to be done by the people directly involved in the production, management, and use of scientific inquiry on specific policy problems. Ideally, most major inquiries would include several such practical experiments in the design of critical fora. Some provision should obviously be made for the periodic comparison of results among programs.

Finally, it should be evident that no innovations in critical criteria or institutions will do much good in the absence of a steady flow of new results to evaluate. In conventional "small" science, this is usually not a problem. With minimal barriers to the entry of new researchers, any new finding or procedure that is sufficiently interesting will soon be played out in a multitude of variations by many independent scholars. But in "big" science, and especially "big" science in policy contexts, the entry barriers are commonly higher. Many scientific inquiries in such contexts tend to be unique ventures built around a single machine, model, or government program. It is crucial to recognize that the "free play of ideas" so necessary to critical progress is severely constrained in such circumstances. However right the answers of single inquiries, their answers are necessarily authoritarian, and their claims to legitimacy are ultimately "numinous," based on being the only game in town. Neither experience nor imagination suggests that a strong, critical consensus on results can possibly develop when a single "official" program, scenario, model, or data set is all that there is to evaluate.

A major challenge for scientific inquiries in policy contexts is therefore to develop means for assuring the "free play of ideas" without which no truly *critical* consensus can exist. This would require a much more pluralistic approach than is now popular to the management and funding of scientific inquiry in policy contexts. At a minimum, explicit mechanisms should be developed for encouraging critical comparisons of the results obtained by independent inquiries on common problems that emerge from the research programs of different nations, government departments, or private institutions. The results of such critical comparisons could only help the strong inquiries, and would serve to remove the manifestly incompetent or biased ones from circulation in a fair and effective manner.

Acknowledgments—We owe a special debt to Jerry Ravetz both for his pathfinding work on the subject of critical appraisal, and for his penetrating and constructive criticism of this manuscript. Welcome counsel and copies of unpublished work were also provided by Alvin Weinberg, Brian Wynne, Michael Thompson, Helga Nowotny, and Steve Rayner. Lisa Carrol contributed greatly to whatever style and clarity the paper may have. Part of this work was supported by the U.S. Department of Energy.

Notes

1. See, for example, the recent discussion of "Linking Science to Policy" in a special issue of *Science, Technology, & Human Values*, Volume 9, Number 1 (Winter 1984).
2. Charles E. Lindblom and David K. Cohen, *Usable Knowledge: Social Problem Solving* (New Haven, CT: Yale University Press, 1979).
3. Stephen Toulmin, *Human Understanding: The Collective Use and Evolution of Concepts* (Princeton, NJ: Princeton University Press, 1972), p. 231; the embedded quote is from W. Lawrence, *Lectures on Physiology, Zoology and the Natural History of Man*, (London, England: Kaygill edition, 1882), p. 95.
4. For examples of this genre, see Martin Greenberger, Mathew A. Crenson, and Brian L. Crissy, *Models in the Policy Process* (New York: Russell Sage Foundation, 1976); David Berlinski, *On Systems Analysis* (Cambridge, MA: The MIT Press, 1976); Harvey M. Sapolsky, *The Polaris System Development* (Cambridge, MA: Harvard University Press, 1972); G. Brewer, *The Politician, the Bureaucrat and the Consultant*, (New York: Basic Books, 1973); William C. Clark, Dixon D. Jones, and C.S. Holling, "Ecological Policy Design," *Ecological Modeling*, Volume 7 (1979): 1–53.
5. For example, Carol H. Weiss, *Evaluating Action Programs* (Boston, MA: Allyn and Bacon, 1972); Carol H. Weiss, *Evaluation Research: Method of Assessing Program Effectiveness* (Englewood Cliffs, NJ: Prentice-Hall, 1972); Daniel P. Moynihan, *Maximum Feasible Misunderstanding* (New York: The Free Press, 1969); Clark Abt, ed., *The Evaluation of Social Programs* (Beverly Hills, CA: Sage Publications, 1977); Jeffrey L. Pressman and Aaron B. Wildavsky, *Implementation* (Berkeley, CA: University of California Press, 1973).
6. See, for example, Karl R. Popper, *The Open Society and its Enemies*, 2 volumes (Princeton, NJ: Princeton University Press, 1962); Ida R. Hoos, *Systems Analysis in Public Policy: A Critique* (Berkeley, CA: University of California Press, 1972); Robert Lilienfield, *The Rise of Systems Theory: An Ideological Analysis* (London, England: John Wiley & Sons, 1978); Jerome R. Ravetz, *Scientific Knowledge and its Social Problems*, (Oxford, England: Oxford University Press, 1971).
7. Joint Economic Committee, U.S. Congress, *Adequacy of U.S. Oil and Gas Reserves* (Washington, DC: U.S. Government Printing Office, 1975).
8. *Ibid.*, p. 53. This and other quotations from the committee hearings are taken from Aaron Wildavsky and Ellen Tennenbaum, *The Politics of Mistrust: Estimating American Oil and Gas Resources* (Beverly Hills, CA: Sage Publications, 1981).
9. Joint Economic Committee, *op. cit.*, p. 14.
10. Wildavsky and Tennenbaum, *op. cit.*, p. 235.
11. Carol H. Weiss, "Policy Research in the Context of Diffuse Decision Making," *Policy Studies Review Annual*, Volume 6 (1982): 19–36.
12. Wolf Haefele, "Hypotheticality and the New Challenges: The Pathfinder Role of Nuclear Energy," *Minerva*, Volume 109 (1974): 303–323.
13. Alvin M. Weinberg, "Science and Trans-science," *Minerva*, Volume 10 (1972): 209–222.

14. See, for example, the National Research Council's agonized discussion of the statistical inference problems involved in detecting the climate signal predicted by general circulation models of the greenhouse effect, in NRC, Carbon Dioxide Assessment Committee, *Changing Climate* (Washington, DC: National Academy of Sciences, 1983), Chapter 5.
15. Oscar Handlin, *Truth in History* (Cambridge, MA: Harvard University Press, 1979), p. 408. This critical stance emphasizes what Handlin goes on to call "the integrity of the record . . . [and] a vital difference in tolerance between facts and interpretation. . . . The correct date, the precise phrasing, the seal were facts which might present difficulties of verification, but which nevertheless admitted of answers that were right or wrong. On the other hand, discussion of opinions and meanings often called for tolerance among diverse points of view, tolerance possible so long as disputants distinguished interpretation from the fact, from the thing in itself. Scholars could disagree on large matters of interpretation; they had a common interest in agreeing on the small ones of fact which provided them grounds for peaceful discourse."
16. Ravetz, *Scientific Knowledge, op. cit.*
17. *Ibid.*
18. E.S. Quade, *Analysis for Public Decisions* (New York: Elsevier, 1975), Chapter 20; Giandomenico Majone and E.S. Quade, eds., *Pitfalls of Analysis* (Chichester, England: John Wiley & Sons, 1980).
19. Edward Shils, ed., *Criteria for Scientific Development* (Cambridge, MA: The MIT Press, 1968).
20. Alvin M. Weinberg, "Criteria for Scientific Choice," *Minerva*, Volume 1 (1963): 159–171; Alvin M. Weinberg, "Criteria for Scientific Choice II: The Two Cultures," *Minerva*, Volume 3 (1964): 3–14; Ravetz, *op. cit.*, pp. 159–169.
21. Toulmin, *op. cit.*, Chapters 2, 3, and 6.
22. Jerome R. Ravetz, "Uncertainty, Ignorance, and Policy." Paper presented at the International Forum on Science and Policy (Laxenburg, Austria: International Institute for Applied Systems Analysis, January 1984).
23. Jerome R. Ravetz, *Scientific Knowledge, op. cit.*, 167.
24. Carol H. Weiss, "Evaluation Research in the Political Context," Volume 1, pp. 13–25, in E.S. Struening and M. Guttentag, eds., *Handbook of Evaluation Research* (London, England: Sage Publications, 1975).
25. Amitai Etzioni, "Two Approaches to Organizational Analysis: A Critique and a Suggestion," *Administrative Science Quarterly*, Volume 5 (1960): 257–278, p. 258.
26. Thomas S. Kuhn, *The Structure of Scientific Revolutions*, second edition (Chicago, IL: University of Chicago Press, 1970).
27. Imre Lakatos, "History of Science and its Rational Reconstructions", in Roger C. Buck and Robert S. Cohen, eds., *Boston Studies in the Philosophy of Science*, Volume 8 (1971): 91–235.
28. Imre Lakatos, "Falsification and the Methodology of Scientific Research Programmes," in Imre Lakatos and Alan Musgrave, eds., *Criticism and the Growth of Knowledge* (Cambridge, England: Cambridge University Press, 1970).
29. Hylton Boothroyd, *Articulate Intervention* (London, England: Taylor and Francis, 1978); Giandomenico Majone, "Policies as Theories," *Omega*, Volume 8 (1980): 151–162.
30. Majone, *op. cit.*, p. 158.
31. *Ibid.*, p. 159.
32. Seymour M. Lipset, *Political Man: The Social Bases of Politics* (Garden City, NY: Anchor Books, 1963), p. 79.
33. Boothroyd, *op. cit.*
34. C.E. Lindbolm, *The Policy Making Process*, second edition (Englewood Cliffs, NJ: Prentice-Hall, 1980), p. 12.
35. Max Weber, *The Theory of Social and Economic Organization* (New York: Free Press, 1922; 1957).
36. Harvey Brooks, "Expertise and Politics: Problems and Tensions," *Proceedings of the American Philosophical Society*, Volume 119 (1975): 257–261, 257.
37. Weinberg, "Science and Trans-science," *op. cit.*
38. Guglielmo Ferrero, *The Principles of Power: The Great Political Crises of History* (New York: Putnam, 1942).
39. Lakatos, "Falsification," *op. cit.*
40. Giandomenico Majone, "Process and Outcome in Regulatory Decision Making," *American Behavioral Scientist*, Volume 22 (1979): 561–583.
41. Marc J. Roberts, Stephen R. Thomas, and Michael J. Dowling, "Mapping Scientific Disputes That Affect Public Policymaking," *Science, Technology, & Human Values*, Volume 9, Number 1 (Winter 1984): 112–122.
42. Brian Campbell, "Uncertainty in Science and Policy," Unpublished Manuscript (Lancaster, England: University of Lancaster, 1983).