Oxford Handbooks Online

Environmental Policy Evaluation and the Prospects for Public Learning

Lawrence Susskind and Alexis Schulman

The Oxford Handbook of U.S. Environmental Policy
Edited by Michael E. Kraft and Sheldon Kamieniecki

Print Publication Date: Oct 2012 Subject: Political Science, U.S. Politics
Online Publication Date: Jan 2013 DOI: 10.1093/oxfordhb/9780199744671.013.0030

Abstract and Keywords

This article reviews conventional approaches to environmental policy evaluation, outlines their presumed relevance to policy making and implementation, and points out the main reasons why they have been subject to challenge. It contrasts the conventional approach to environmental policy evaluation—which presumes the identity of the policy analyst is unimportant—with the "collaborative approach," which emphasizes the need to engage relevant stakeholders (i.e., the users of policy analyses and those affected by them) in the process of environmental policy evaluation. The article also describes the emergence of "adaptive" approaches to resource management and sustainable development, and explains why they represent an important shift away from emphasizing "success" and "failure" in environmental policy making and toward ongoing public learning for purposes of improvement.

Keywords: environmental policy, policy making, collaborative approach, adaptive approach, resource management

1. Introduction

In most models of public policy making, the evaluation stage comes after implementation, but before reconsideration of goals, objectives, and policy designs. Evaluation is supposed to facilitate improvements and corrections. The tools and techniques of policy evaluation have received a great deal of attention in the public policy literature, including quantitative tools like cost-benefit analysis and multivariate analysis of large data sets, as well as qualitative tools such as case studies and ethnographic accounts of the results of new programs or policy ideas (Cook and Campbell 1979; Patton 1980; Stokey and Zeckhauser 1978; Yin 1994, 2002). Most accounts of environmental policy making and implementation take the same approach to evaluation: (1) it is in the public interest to determine whether environmental policies, at every level, and the programs designed to implement them, have worked; (2) this can be accomplished by studying the intended and

unintended effects of environmental policies using appropriate social science research tools; (p. 678) and (3) the results of such studies should allow us to "do better" the next time around.

This presumes a number of things. First, it presumes that environmental policy analysts can sort out the effects or impacts of new environmental policies and programs from the effects of everything else that has happened at the same time. That is a heroic assumption. Second, it presumes that these same policy analysts will be able to tally the impacts or effects of the environmental policy they want to study, in light of some agreed-upon benchmarks. This will make clear whether the results were "worth it." Third, it presumes that these studies will allow policy makers to see clearly what they ought to do differently the next time around. All three presumptions have been strongly contested (Fischer 1980, 1995; Ingram and Mann 1980; Manzer 1984; Packwood 2002).

This chapter reviews conventional approaches to environmental policy evaluation, outlines their presumed relevance to policy making and implementation, and points out the main reasons why they have been subject to challenge. Then we contrast the conventional approach to environmental policy evaluation—which presumes the identity of the policy analyst is unimportant—with what we call the "collaborative approach," which emphasizes the need to engage relevant stakeholders (i.e., the users of policy analyses and those affected by them) in the process of environmental policy evaluation. We also describe the emergence of "adaptive" approaches to resource management and sustainable development and explain why they represent an important shift away from emphasizing "success" and "failure" in environmental policy making and toward ongoing public learning for purposes of improvement. "Collaborative adaptive management," the name now given to this new approach, is one of most important developments in environmental policy making. It represents a shift away from the assumption that our scientific understanding of complex socio-ecological systems is sufficient to justify setting precise standards and long-term goals. Rather, it assumes that the complexity and uncertainty surrounding the search for sustainable ways of managing resources and directing growth require us to move more slowly and carefully, with an emphasis on continuous monitoring and step-by-step adjustment.

2. The Domains of Evaluation

Most of the time, policy evaluation has been focused on whether particular environmental policies achieved their intended goal(s). For example, did the federal government's efforts to reduce lead levels in ambient air achieve that result? We call this the first domain of evaluation. Analysts doing policy evaluation in this domain try to measure the impacts or results of policies and programs in light of what was intended. This is not, however, the only way to frame a policy or program evaluation. For example, analysts working in what we call the second (p. 679) domain of evaluation ask, "What positive and negative impacts, in the long and short term, did lead-reduction levels have, regardless of what was intended?" They focus on unintended consequences, which sometimes turn out to be very

important. Analysts working in the third domain of evaluation ask, “In light of what we now know, do the original goals or objectives of federal lead-reduction policies still make sense?” Evaluation in the third domain questions the original intent or objectives. As time goes on, it may turn out that even policies and programs that were entirely successful no longer merit support. Finally, evaluators working in the fourth domain ask, “What other policies or policy approaches, advanced in different ways, might be more beneficial (and to whom) than the policy or program we were initially asked to evaluate?” This requires comparing something that has happened with alternatives that have not yet been tried, or at least not implemented in the same locale(s) in the same time period. The tools and techniques needed to answer these four kinds of questions, and the problems they raise, are quite different (Bardach 2012; Haas and Springer 1998; Hogwood and Gunn 1984; Weimer and Vining 2011).

Most environmental policy evaluation begins in the first domain. Congress, state legislators, and city councils want to know whether the laws and regulations they enacted are working. Very quickly, though, evaluation in the first domain leads curious or self-interested researchers to ask second-, third-, and fourth-domain questions. Were there unintended effects? Are there other ways to achieve the same objectives that might be more effective? Are the original policy objectives still relevant? It is difficult to cumulate research on the effectiveness of an environmental policy if all the available evaluative studies begin with different questions and use different methods to answer them.

3. The Conventional Approach to Public Policy Evaluation

By the 1960s and 1970s, systematic and empirically grounded evaluation and research were widely viewed as the “proper basis for decision-making in public policy” (Fischer 1995, 4). For the first time, analysts were directly responsible for the development of public programs, such as those spearheaded under Lyndon Johnson’s “Great Society”; public agencies were retooled to expand their research capacities; and by the 1970s, Congress had turned its sights on evaluation, creating the Congressional Budget Office and Office of Technology Assessment and promoting the use of evaluation in a suite of laws (Fischer 1995). Policy professionals were optimistic that by applying quantitative tools and “scientific” methods they would be able to figure out what was working, what needed improvement, and what else might be tried. **(p. 680)**

However, as the field evolved, it ended up adopting a limited perspective, confining itself to the task of answering questions in the first domain—narrowly defined actual or expected empirical outcomes of given policy goals (Fischer 1995). This focus was coupled with a rigid “scientific approach,” which assumed that any qualified expert using prescribed (quantitative) methods would get replicable results. Criticism of this rationalist, expert-guided approach gave rise to other (more relativistic) schools of evaluative theory and practice, variously entitled “postpositivist” or “argumentative/deliberative.”[1] However, a great deal of environmental policy evaluation, whether before-

the-fact assessments of policies or programs, or after-the-fact analyses of outcomes, still remains grounded in the positivist optimism of the 1960s and 1970s.

Contemporary policy evaluations are undertaken for many reasons by various administrative, legislative, and advocacy organizations. They may be required by law, requested by decision makers, initiated by various stakeholders or demanded by the public at large. They can be completed by a range of actors in different institutional settings, from think tanks, to academic research centers to political oversight bodies. Even given these different starting points, most environmental policy evaluations in the public arena follow a general six-step process:

1. Define the issue and suggest a method of evaluation.
2. Choose a consultant or a study team.
3. Write a contract spelling out the obligations of the evaluators.
4. Perform the evaluation.
5. Submit a draft report for comment.
6. Produce and disseminate findings and recommendations. (Susskind, Jain, and Martyniuk 2001, 10)

This approach presumes that these six steps can and should involve a purely rational process, separating fact (analysis) from values (politics). The feasibility of this enterprise, however, rests on several key assumptions (Bovens, Hart, and Kuipers 2006). It requires (1) that the goals and objectives of particular environmental policies are clear at the outset; (2) that there is agreement on which indicators should be used to gauge policy outputs and outcomes; and (3) that the views and loyalties of the analyst (aside from their technical competence) should have no impact on the outcome, as long as the right methods are used in the right way. However, the goals of specific environmental policies are almost always murky, in part because of the need to satisfy a sufficient number of stakeholders to win political support. And the choice of evaluative criteria, or benchmarks of success, is contested as part of ongoing disciplinary and ideological battles. What gets measured is based on underlying causal models favored by some researchers but rejected by others.

Thus, in practice, there are too many nonobjective judgments that must be made along the way for the outcome not to reflect, at least in part, the biases or (p. 681) ideological predispositions of the evaluators and their clients. Value-laden decisions must be made at each step in the process. For example, study sponsors, more often than not, have a specific desired outcome in mind. By hiring a consultant who shares the sponsor’s bias and defining the scope of their evaluation (steps 1–3), it is possible for sponsors to obtain evaluation results that are largely preordained (Susskind, Jain, and Martyniuk 2001).

Furthermore, the decisions and trade-offs made during analysis (step 4), regardless of the methodological approach selected, have a major impact on findings. Quantitative approaches, seeking to simulate controlled experimental designs, dominate the field. The benefits of these methods are well recognized. They allow analysts to draw generalizations about populations from smaller, statistically chosen samples, and to

establish the relationship of chosen variables to policy outcomes, and they make it possible to speculate on the relationship between causes and effects. In theory, the findings of such evaluations are replicable, meaning they should not hinge on the background of the evaluator.

Since the 1980s, qualitative methods, like case studies and participant observation, have gained in popularity in the environmental policy field because they provide unique, if not complementary, policy knowledge (Sadovnik 2007; Yanow 2007). While these methods may be weak where quantitative methods are strong—generalizability, verification, and reliability—they can enrich our understanding of how environmental policies are working by providing detailed accounts that pay attention to context, revealing why and how certain policy outcomes may have occurred (Sadovnik 2007). Since policy makers and stakeholders may be more interested in causation than correlation, qualitative analyses may be the most effective means of enabling evaluators to make convincing causal arguments.

Although proponents of quantitative approaches have long argued that their methods are inherently more objective than qualitative ones, requiring less interpretive work (Campbell and Stanley 1963; Page and Stake 1979), both hinge on nonobjective judgments. At various points in a quantitative analysis, for example, it is necessary to limit the time and geographic scope of an investigation, choose indicators and set discount rates, and integrate or sum up across multiple dimensions of analysis. As we discuss further in the following sections of this chapter, there are no objectively correct ways of doing these things. Resource or time limitations increase the likelihood that evaluators will make arbitrary choices that affect their findings.

Finally, interpreting study findings, drawing conclusions, and making recommendations (steps 5 and 6) involve what has been called the “normative leap.” Moving from what has been found to what ought to be done requires evaluators to make still more nonobjective judgments (Schön and Rein 1994). At the point where conclusions and recommendations must be drawn, study sponsors typically exercise significant discretion, “softening” findings that do not comport with their expectations, reframing conclusions, and deciding what to make public.

This inherently political and value-laden process “virtually guarantees controversy,” particularly when multiple evaluations are commissioned of the same (p. 682) policy initiative (Susskind, Jain, and Martyniuk 2001, 7). Not surprisingly, evaluators drawing on different methods and findings rarely come to the same conclusion, leaving policy makers no better positioned to make a wise decision:

> In the ideal world of the positivist social scientist, we stand to gain from this multiplicity: presumably it results in more facts getting on the table.…In the real world, multiple evaluations of the same policy tend to be non-cumulative and non-complementary. (Bovens, Hart, and Kuipers 2006, 321)

The obstacles facing decision makers who make use of environmental policy analyses are not limited to having to reconcile contradictory findings. In an effort to understand why so few evaluations actually produce better policy results and are instead ignored by decision makers (Webber 1992; Weiss 1979), researchers have zeroed in on the dynamics of how knowledge is received and disseminated (Kraft 1998). Knott and Wildavsky (1980) proposed "seven standards of utilization" that affect how the results of policy evaluations are interpreted:

1. *Reception*, which occurs when results reach the decision maker.
2. *Cogitation*, which occurs when the decision maker reads, digests, and otherwise thinks about the findings.
3. *Reference*, which occurs if the study has somehow changed the decision maker's preference or worldview.
4. *Effort*, which occurs if and when the study influences the action of the decision maker.
5. *Adoption*, which occurs when the study actually influences policy outcomes.
6. *Implementation*, which occurs when adopted policy becomes practice.
7. *Impact*, which occurs when tangible benefits to society have been realized.

Although the first standard is usually met, there are a number of factors that routinely impede the achievement of the others. First*, cogitation*, *reference*, and *effort*, the standards most influenced by the characteristics of environmental policy evaluations as opposed to the determination and resources of the decision maker, are unlikely to be met if the wrong questions were asked—that is, if the policy evaluation framed the issue in an unhelpful manner, given the concerns of the decision maker. Second, if the study results are not understandable because of the complexity of the methods used, for example, the chances that a decision maker will actually read and comprehend a study are significantly reduced. Third, if the findings of an evaluation are not convincing, providing a new frame of reference, the study is unlikely to meet standards 5–7 (adoption, implementation, and impact). Whether this is the case is a product of the reputation and credibility of the analysts and sponsor (for example, is there a clear bias, and was the study politicized?) and the manner in which the study was conducted.

The conventional approach to environmental policy evaluation frequently fails to produce evaluations that meet all the standards for effective knowledge (p. 683) utilization listed above. The unfortunate result is that while many analyses are conducted, few actually produce wiser policy decisions.

4. Conventional Approaches to Policy Evaluation Are Not Well Suited to the Environmental Policy Field

Because of the complexity of the socio-ecological systems involved, the conventional ways of gauging the success of public policies and programs are not entirely applicable in the environmental policy field. The usual assumptions about measuring costs and benefits create special difficulties for environmental policy evaluators. It is hard to know how to conceptualize gains and losses, for example, when we are trying to study the effectiveness of policies aimed at protecting endangered species or restoring damaged ecosystems. The amount of money saved or spent is hardly relevant to the primary goal of ensuring that natural systems are functioning properly.

Although many environmental policy evaluations do incorporate the monetary valuation of costs and benefits, multiple values are almost always at stake. For example, issues of endangered species conservation invariably involve concerns that go well beyond utilitarian considerations. Environmental issues are ineluctably linked to questions of equity, governance, and even spirituality. The multidimensionality of environmental issues means that a “plurality of legitimate perspectives” can surround the same problem (Funtowicz and Ravetz 1993, 739). Although the core values, interests, and assumptions that result from these disparate positions are often taken for granted, they really ought to be factored into environmental policy evaluation. So, efforts to rely on a unitary measure of value are not only misleading, but also bound to incite conflict.

Further complicating efforts at environmental policy evaluation is the matter of data. For many environmental processes, there are large data gaps. In some instances, operational criteria are not clear and there are no records of continuous performance—all of which are crucial to ex-post evaluation. For example, while solid waste management at the local level is generally well monitored by government agencies, data on recycling streams are still scarce (Crabbé and Leroy 2008). In some instances data exist but cannot be shared because of proprietary or legal restrictions. Even when relevant data can be located, inconsistencies in terminology, collection methods, and questions of reliability present further challenges (Solomon 1998). Data overload can be just as problematic. Mining for relevant data, patching together data sets, or performing comparative analyses with incongruent sources is more of an art than a science, and leaves considerable discretion to the analyst. **(p. 684)**

Even with high-quality data, the complexity of the socio-ecological systems—composed of a vast number of interacting variables through both space and time—makes it particularly difficult to model the dynamics involved. Without reliable models, it is hard to analyze whether particular policy interventions will or have produced the desired results. Setting the time frames for environmental policy evaluation is also problematic, raising questions about when we can reliably see intended results or judge a policy to be successful. For example, a great many environmental impact assessments of proposed new hydroelectric

plants failed to consider the long-term buildup of mercury in aquatic life and the effects of such changes on the well-being of children (Rosenberg, Bodaly, and Usher 1995).

This system complexity also embeds a great deal of uncertainty in environmental policy evaluations. Policy evaluators attempting to assess the desirability of various environmental policy options before or after the fact face uncertainty about the likelihood that particular outcomes will occur, as well as uncertainty regarding the actual characteristics of the outcomes that have transpired. In environmental policy analysis, there are tools like risk assessment, scenario analysis, and sensitivity analysis that can be used to hedge against uncertainty; however, in the environmental policy evaluation field there are no agreed-upon rules about how these techniques, and others like them, should be used (meaning that they are interpreted quite differently by individual evaluators).

To cope with data gaps, system complexity, and the resulting uncertainty, evaluators must make numerous simplifying assumptions. These are inherently political and subjective. For example, no scientific procedure specifies, a priori, a problem definition or a proper scope of analysis. Setting an analytical frame—the description of the system, how it should be bounded in space and time, and what variables should be considered important —introduces all kinds of bias, particularly epistemological bias (Munda 2000). Indeed, different problem framing can yield vastly different results. For example, between 1979 and 2003, evaluations of the environmental costs of new coal power plants varied by a factor of 50,000— a result of the way scientific models were developed and integrated into the evaluation process (Stirling 1997, as cited in Munda 2000).

5. Collaborative Approaches to Environmental Policy Evaluation

One way, and perhaps the best way, of defending the subjective judgments that are an inevitable part of environmental policy evaluation is to adopt a collaborative or participatory approach to making such assumptions. That is, by involving the relevant stakeholders or their representatives in reviewing the nonobjective judgments essential to environmental policy evaluation (what indicators of impact to (p. 685) use, what geographic scope to set for an evaluative study, or what time frame to specify), it is possible to avoid many, if not all, of the challenges that undermine the political credibility and stymie the use of environmental policy evaluations done in the conventional way. Frequently bundled with “participatory policy analysis” in the evaluation literature, the collaborative approach can be understood as both a strategy for dealing with conflict and a “societal response to changing conditions in increasingly networked societies…where differences in knowledge and values among individuals and communities is growing” (Innes and Booher 1999, 412). Collaborative approaches to environmental policy evaluation are particularly well suited to the type of “socio-technical” (Fischer 1995, 222) or “wicked”[2] problems characteristic of environmental policy dilemmas.

Collaborative evaluation shares a common history with other types of postpositive approaches to public decision making. All were developed in response to perceived flaws in positivist, expert-driven methodologies. They reject the fact/value dichotomy inherent in rationalist evaluation. Postpositivists view evaluation as the continuation of politics by "other means" and note how bias is mobilized even within its initial stages. As Bovens, Hart, and Kuipers (2006, 327) write, "Evaluation simply mirrors the front end of the policy process (agenda setting and problem definition): some group's interests and voices are organized 'in,' whereas other stakeholders are organized 'out.'" The mission of evaluation, then, is not to provide the one best policy choice or final judgment but, rather, to develop knowledge and come to decisions through argumentation among apparently contradictory perspectives (Bovens, Hart, and Kuipers 2006; Fischer 1998). Drawing on the work of deliberative democrats and theories of communicative rationality,[3] the "argumentative" turn in evaluation cites the transformative role of rational discourse. As Innes and Booher explain, "The basic idea of communicative rationality is that emancipatory knowledge can be achieved through dialogue that engages all those with differing interests around a task or problem" (1999, 6).

Collaborative evaluation embraces the theory of communicative rationality, but elaborates a specific *process* of knowledge development and decision making that is drawn from the fields of consensus building and dispute resolution. This approach diverges from other participatory evaluation methods, such as deliberative polling (Fishkin 1991), in that the goal is to build an informed consensus by engaging all relevant stakeholders—representing both majority and minority interests—in joint decision making. By involving all interested parties (including decision makers, analysts, and the stakeholding public) in a fair and transparent way, environmental policy evaluations prepared in this fashion are far more likely to minimize the conflict that surrounds "expert" assessments, increase the perceived legitimacy of the outcome, and produce salient assessments that facilitate (p. 686) public learning and wise decisions (Innes 1999; Scher 1999; Susskind, Jain, and Martyniuk 2001).

Although there are many ways to undertake a collaborative environmental policy evaluation, there are three key conditions that must be met. First, any group that believes it is a stakeholder should be consulted about the design of the process and the selection of actual participants. Second, a professional facilitator who does not represent the sponsoring agency or organization should be selected to manage the work. Last, all participants in the process should provide written comments on any interim products of any evaluation. The following elaborates further on the basic steps in collaborative evaluation: convening; assigning roles and responsibilities; facilitating group problem solving; reaching agreement; and holding parties to their commitments (Susskind, McKearnan, and Thomas-Larmer 1999).

A collaborative assessment process needs to be convened by someone with the authority to do so. Once a convener, usually a public agency, decides to proceed, the first steps in a collaborative evaluation include identifying the relevant stakeholders and assessing their

concerns. When a decision is made to begin a policy evaluation formally, roles and responsibilities are assigned.

The bulk of a collaborative assessment process is typically spent in group work. How this is undertaken is unique to each group. However, the basic principles include these: strive for transparency; seek expert input when engaging in joint fact-finding (see below for more details); create working subcommittees if appropriate; use the help of a skilled facilitator; and use a single-text procedure.

The final stages of a collaborative evaluation include seeking unanimity on a final report, checking back with each participant's constituents, and producing a written record of the understandings that have been reached (Susskind and Cruikshank 2006).

Given the technical nature of most environmental policy issues, joint fact-finding (JFF) may be an integral part, or even primary focus, of a collaborative policy evaluation. JFF is a procedure for involving those affected by policy decisions in the process of generating and analyzing the scientific and technical information that will be used to inform value-laden decisions. The goal is to produce scientific and technical information that is salient, credible, and legitimate through a process that produces shared learning, trust, creative problem solving, and shared ownership (Erhmann and Stinson 1999). As noted, JFF can be embedded in a larger evaluation process, particularly when there is disagreement over the scientific and technical information that ought to be included. The principles of JFF are similar to those of consensus building more generally. That is, stakeholder groups should be fully represented and self-selected; stakeholders should specify the research protocol together (including technical advisors); and a neutral facilitator is usually required to help the group reach a formal, written agreement.

JFF brings stakeholders together with technical experts to refine the questions that will be asked and select the most appropriate methods of answering them. The participants in JFF assess information and data gaps, reframe general questions as specific ones, determine the strengths and shortcomings of various analytic (p. 687) tools, and gather and analyze chosen data. Once an analysis is complete, the group comes to an agreement on how to use the results—how to deal with conflicting interpretations and how to clarify remaining uncertainties and possible contingent responses. Finally, before results are communicated to constituencies and decision makers, stakeholders jointly review final drafts and studies, determine whether further JFF is necessary, and integrate findings into final recommendations.

JFF has been used successfully to evaluate environmental policy options in complex and often volatile political contexts. For example, such a process was used to generate policy recommendations on how water delivery commitments along the Columbia River basin could be met with minimum impact on Glen Canyon National Recreation Area and Grand Canyon National Park. Stakeholders and experts jointly scoped and conducted research with the help of a neutral facilitator and reached consensus on a "single text" that included findings and recommendations (Council on Environmental Quality 2007). JFF may take more time, cost a bit more, and put policy evaluators in what, for some, will be

the uncomfortable position of having to interact with nonexperts. But when recent evaluation efforts have been unsuccessful and policy makers want to be sure that "local knowledge" is incorporated into subsequent rounds of decision making, JFF—and more collaborative assessments in general—make sense.

6. Adaptive Management: A Good Reason for Evaluation

While collaboration, including joint fact-finding, is an approach to ensuring transparency and enhancing the legitimacy and political credibility of environmental evaluation studies, collaboration in and of itself will not ensure good results. In our view, collaborative approaches to environmental policy evaluation need to be tied to an adaptive approach to environmental management to increase the chances that an evaluation will have a significant impact on decision making (Susskind, Camacho, and Schenk 2010). Given the complexity of environmental systems, and the uncertainties that surround most socio-ecological interactions, almost all environmental policy making is likely to be off the mark, at least at the outset. It makes more sense to think of environmental evaluation as a means of supporting ongoing adjustments or what might be called public or social learning.

From the beginning of each environmental-policy-making effort, attention should be paid to how implementation will be monitored, who will have responsibility for making sense of the findings, and how both ends and means can be continuously adjusted in light of what is learned.

Adaptive management imagines evaluation as something that happens during rather than after policy objectives have been set. For example, objectives like (p. 688) the achievement of "fishable" or "swimmable" water, reductions in greenhouse gas emissions, or the restoration of contaminated areas to their "original" state are general goals that typically need to be recalibrated as more is learned—in order to reshape ongoing programmatic efforts.

7. The Dynamics of Public Learning

For the results of environmental policy evaluations to have an impact on public decision making, some degree of learning is necessary. "Perfunctory" or "technical" learning is, unfortunately, the most common. Because it seeks only to validate existing policy or consider alternative means within the same goal structure, it is not likely to lead to substantial improvement (Howlett, Ramesh, and Perl 2009). "Contested" learning or social learning that seeks to harmonize competing evaluations by multiple stakeholders and confronts radically different ways of doing things is more likely to lead to marked improvements in environmental outcomes, but this is much less common.

It is quite possible for “nonlearning” (i.e., no learning at all) and other forms of “limited learning” to occur in environmental-policy-making and environmental management situations. Nonlearning involves a failure to undertake a serious review of the results when specific means are used to achieve various desired ends. Limited learning occurs when lessons of only a very restricted sort are drawn during the evaluation process (Howlett, Ramesh, and Perl 2009). Moreover, what individuals learn is one thing; what organizations or networks learn—social learning—is something quite different. The more “open” a policy network is, the greater the extent to which a wide range of stakeholders will have a chance to participate fully in an assessment process. The more fully a wide range of stakeholders participates, the more likely it is that social learning can occur. Of course, when a broader set of assessment techniques is embraced by policy actors with a wide range of viewpoints, there is a possibility that “contestation” will lead to nothing more than stalemate. Contested learning is more likely to produce social learning and not stalemate when a collaborative approach is used to facilitate the emergence of an informed consensus.

Thus, there are two key obstacles to social learning. One is the willingness and capacity of government employees to engage in joint fact-finding and collaborative assessment with a wide range of stakeholders who may be critical of existing policies and programs. The second is the “the nature of the policy sub-system,” that is, the extent to which it is open or closed (Howlett, Ramesh, and Perl 2009). In open subsystems, stakeholder elites and government officials are willing to consider modifications in what they are doing (if only to prove their worth). In closed subsystems they are not. Only if these obstacles to social learning can be removed, usually through the leadership of a few key individuals, will the promise of adaptive (p. 689) management, and the challenge to conventional notions of environmental policy evaluation it represents, be realized.

8. Conclusions

The evolution of environmental policy evaluation should be viewed within the larger context of public policy making or policy science, relatively new fields with a strong commitment to rational analysis. Unfortunately, the unique aspects of environmental systems, particularly their complexity and the uncertainty it creates, make the conventional approach to public policy evaluation particularly difficult in the environmental field. Almost all environmental policy evaluation begins in the first domain, focusing on whether particular targets, such as ensuring “fishable” or “swimmable” water quality, reducing greenhouse gas emissions to pre-1992 levels, or restoring contaminated rivers to their “original” state have been met. However, the unintended or second-order effects soon attract attention. For example, how have attempts to achieve these targets shortchanged other important goals? Then, too, radically different ways of achieving the same objectives—by using new pricing strategies rather than mandating the use of best available technologies—emerge and there is pressure to figure out which will be the most cost-effective or sustainable strategy. Eventually, someone asks whether we really need to keep worrying about these objectives, or whether different problems

deserve our attention. The work of environmental policy evaluators in all four domains requires them to make a great many nonobjective judgments so they can delimit what they are evaluating. The only way these judgments, and the results they produce, can be credible is if the relevant stakeholders are involved in specifying what questions need to be answered, which methods of evaluation should be used, how gaps in the data and uncertainties of various kinds should be handled, and how the findings should be interpreted. Such a collaborative approach makes sense, especially in a context in which incremental approaches to management—that stress monitoring and careful adjustments in both means and ends—seem like the most sensible way of proceeding given our inability to model the complexities of the natural-societal systems involved. Incremental approaches, like adaptive management, assume that environmental policies and programs need to be adjusted continuously, creating a demand for social learning. This is an approach to environmental policy evaluation that puts a premium on how to do better, rather than on what has succeeded and who has failed.

Environmental policy analysis, as a field, depends on careful reviews of the ways in which specific evaluation studies are used, or not, in actual policy making. It would help if one or more research centers dedicated its efforts to compiling such reviews and offering periodic summaries of its findings. This could be done online by creating a “wiki” that encourages environmental policy evaluators to report their findings in a consistent format. Those who track the impact of evaluations, done in different (p. 690) ways, on decision making should make special efforts to highlight the assumptions that evaluators are making when nonobjective judgments come into play. Finally, in terms of a substantive research agenda, it would be helpful to know more about the added value that collaboration in environmental policy evaluation yields.

References

Bardach, E. 2012. *A Practical Guide for Policy Analysis: The Eightfold Path to More Effective Problem Solving*. 4th ed. Washington, DC: CQ Press.

Bovens, M., P. T. Hart, and S. Kuipers. 2006. “The Politics of Policy Evaluation.” In *The Oxford Handbook of Public Policy*, ed. M. Moran, M. Rein and R. E. Goodin. New York: Oxford University Press.

Campbell, D. T., and J. Stanley. 1963. *Experimental and Quasi-experimental Designs for Research*. Boston: Houghton Mifflin.

Cook, T. D., and D. T. Campbell. 1979. *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally College.

Council on Environmental Quality. 2007. *Collaboration in NEPA: A Handbook for NEPA Practitioners*. Available at **ceq.eh.doe.gov/ntf/Collaboration_in_NEPA_Oct_2007.pdf** (accessed January 2011).

Crabbé, A., and P. Leroy. 2008. *The Handbook of Environmental Policy Evaluation*. Sterling, VA: Earthscan.

Erhmann, J. R., and B. L. Stinson. 1999. “Joint Fact-Finding and the Use of Technical Experts.” In *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement*, ed. L. Susskind, S. McKearnan, and J. Thomas-Larmer. Thousand Oaks, CA: Sage.

Fischer, F. 1980. *Politics, Values, and Public Policy: The Problem of Methodology*. Boulder, CO: Westview.

———. 1995. *Evaluating Public Policy*. Chicago: Nelson-Hall.

———. 1998. “Beyond Empiricism: Policy Inquiry in Postpositivist Perspective.” *Policy Studies Journal* 26 (1): 129.

Fishkin, J. S. 1991. *Democracy and Deliberation: New Directions for Democratic Reform*. New Haven, CT: Yale University Press.

Funtowicz, S. O., and J. R. Ravetz. 1993. “Science for the Post-Normal Age.” *Futures* 25 (7): 739–755.

Haas, P. J., and J. F. Springer. 1998. *Applied Policy Research: Concepts and Cases*. New York: Garland.

Habermas, J. 1984. *The Theory of Communicative Action*. London: Heinemann.

Hogwood, B. W., and L. A. Gunn. 1984. *Policy Analysis for the Real World*. Oxford: Oxford University Press.

Howlett, M., M. Ramesh, and A. Perl. 2009. *Studying Public Policy: Policy Cycles and Policy Subsystems*. New York: Oxford University Press.

Ingram, H. M., and D. E. Mann. 1980. *Why Policies Succeed or Fail*. Beverly Hills, CA: Sage.

Innes, J. E. 1999. “Evaluating Consensus Building.” In *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement*, ed. L. Susskind, S. McKearnan, and J. Thomas-Larmer. Thousand Oaks, CA: Sage. (p. 691)

Innes, J., and D. E. Booher. 1999. “Consensus Building and Complex Adaptive Systems: A Framework for Evaluating Collaborative Planning.” *Journal of the American Planning Association* 65 (4): 412–423.

Knott, J., and A. Wildavsky. 1980. “If Dissemination Is the Solution, What Is the Problem?” *Knowledge: Creation, Diffusion, Utilization* 1 (4): 537–578.

Kraft, M. E. 1998. “Using Environmental Program Evaluation: Politics, Knowledge, and Policy Change.” In *Environmental Program Evaluation: A Primer*, ed. G. J. Knaap and J. K. Tschangho. Urbana: University of Illinois Press.

Ludwig, D. 2001. “The Era of Management Is Over.” *Ecosystems* 4 (8): 758–764.

Manzer, R. 1984. “Policy Rationality and Policy Analysis: The Problem of the Choice of Criteria for Decision-Making.” In *Public Policy and Adminstrative Studies*, ed. O. P. Dwivedi. Guelph, ON: University of Guelph.

Munda, G. 2000. *Conceptualising and Responding to Complexity*. Cambridge, UK: Cambridge Research for the Environment.

Packwood, A. 2002. “Evidence-Based Policy: Rhetoric and Reality.” *Social Policy and Society* 1 (3): 267–272.

Page, E. B., and R. E. Stake. 1979. “Counterpoint: Should Educational Evaluation Be More Objective or More Subjective?” *Educational Evaluation and Policy Analysis* 1 (1): 45–47.

Patton, M. Q. 1980. *Qualitative Evaluation Methods*. Beverly Hills, CA: Sage.

Rosenberg, D. M., R. A. Bodaly, and P. J. Usher. 1995. “Environmental and Social Impacts of Large Scale Hydroelectric Development: Who Is Listening?” *Global Environmental Change: Human and Policy Dimensions* 5 (2): 127–148.

Sadovnik, A. R. 2007. “Qualitative Research and Public Policy.” In *Handbook of Public Policy Analysis: Theory, Politics, and Methods*, ed. F. Fischer, G. Miller, and M. S. Sidney. Boca Raton, FL: CRC/Taylor & Francis.

Scher, E. 1999. “Negotiating Superfund Cleanup at the Massachusetts Military Reservation.” In *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement*, ed. L. Susskind, S. McKearnan, and J. Thomas-Larmer. Thousand Oaks, CA: Sage.

Schön, D. A., and M. Rein. 1994. *Frame Reflection: Toward the Resolution of Intractable Policy Controversies*. New York: Basic Books.

Solomon, L. S. 1998. “Evaluation of Environmental Programs: Limitations and Innovation.” In *Environmental Program Evaluation: A Primer*, ed. G. J. Knaap and J. K. Tschangho. Urbana: University of Illinois Press.

Stirling, A. 1997. “Limits to the Value of External Costs.” *Energy Policy* 25 (5): 517.

Stokey, E., and R. Zeckhauser. 1978. *A Primer for Policy Analysis*. New York: W. W. Norton.

Susskind, L., A. E. Camacho, and T. Schenk. 2010. “Collaborative Planning and Adaptive Management in Glen Canyon: A Cautionary Tale.” *Columbia Journal of Environmental Law* 35 (1): 1–56.

Susskind, L., and J. L. Cruikshank. 2006. *Breaking Robert’s Rules: The New Way to Run Your Meeting, Build Consensus, and Get Results*. Oxford: Oxford University Press.

Susskind, L., R. K. Jain, and A. O. Martyniuk. 2001. *Better Environmental Policy Studies: How to Design and Conduct More Effective Analysis*. Washington, DC: Island.

Susskind, L., S. McKearnan, and J. Thomas-Larmer, eds. 1999. *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement*. Thousand Oaks, CA: Sage. (p. 692)

Webber, D. J. 1992. “The Distibution and Use of Policy Knowledge in the Policy Process.” In *Advances in Policy Studies since 1950*, ed. W. N. Dunn and R. M. Kelly. New Brunswick, NJ: Transaction.

Weimer, D. L., and A. R. Vining. 2011. *Policy Analysis: Concepts and Practice*. 5th ed. New York: Pearson Longman.

Weiss, C. H. 1979. “The Many Meanings of Research Utilization.” *Public Administration Review* 39 (5): 426–431.

Yanow, D. 2007. “Qualitative-Interpretive Methods in Policy Research.” In *Handbook of Public Policy Analysis: Theory, Politics, and Methods*, ed. F. Fischer, G. Miller, and M. S. Sidney. Boca Raton, FL: CRC/Taylor & Francis.

Yin, R. K. 2002. *Case Study Research: Design and Methods*. Thousand Oaks, CA: Sage.

Notes:

(1.) For a discussion of these schools of evaluation, see Bovens, Hart, and Kuipers 2006.

(2.) Of wicked problems, Ludwig (2001, 759) writes that they “have no definitive formulation, no stopping rule, and no test for final solution.”

(3.) For the development of this theory, see Habermas 1984.

Lawrence Susskind

Lawrence Susskind is Ford Professor of Urban and Environmental Planning, Massachusetts Institute of Technology, and Vice Chair of the Program on Negotiation, Harvard Law School.

Alexis Schulman

Alexis Schulman is a Ph.D. candidate in the Department of Urban Studies and Planning, Massachusetts Institute of Technology.