

Cooksy, L. J., Mark, M. M., & Trochim, W.M.K. (2009). Evaluation policy and evaluation practice: Where do we go from here? In W.M.K. Trochim, M. M. Mark, & L. J. Cooksy (Eds.), *Evaluation policy and evaluation practice. New Directions for Evaluation*, 123, 103–109.

7

Evaluation Policy and Evaluation Practice: Where Do We Go From Here?

Leslie J. Cooksy, Melvin M. Mark, William M. K. Trochim

Abstract

Three issues for evaluation policy and practice are described: evaluation policy dimensions, evaluation policy instruments, and the political and economic environment for evaluation policy. Selected future directions are outlined, including the need to describe the evaluation policy landscape, further articulate an evaluation policy taxonomy, and develop and implement tactics for influencing evaluation policy, with particular attention to the role of professional associations. © Wiley Periodicals, Inc.

If the context of evaluation practice is largely defined by evaluation policies, then evaluation policy gives us a way to think about systematically influencing that context. Through policy, such principles as technical quality, respect for people, and utility can be explicitly and systematically built into the evaluation expectations of an organization rather than being values we fight for, one evaluation at a time. Of course, it is never this simple (Julnes & Rog, 2007). To be a player in the world of evaluation policy making, we need to develop a language for communicating about types of



policies and strategies for influencing them. In this issue, Trochim lays out a theoretical framework and offers several ways of furthering our knowledge about evaluation policy. Calling for the American Evaluation Association (AEA) to continue its evaluation policy advocacy, Datta gives examples of professional organizations that have influenced evaluation (and other) policies and recommends strategies, such as coordinating with existing consortia of associations, to strengthen AEA's voice. Chelimsky discusses evaluation policies associated with the structure and location of the evaluation function in organizations and describes the kinds of policies needed to protect the independence of evaluators. Stern expands our thinking by examining the issue of evaluation policy from the perspectives and experiences of supranational organizations such as the European Union (EU). Leeuw's historical account of the evolution and increasing formality of evaluation policy in the Netherlands illustrates the role of culture in evaluation policy development.

These authors address various facets of evaluation policy and consider those facets in different settings, but plenty of territory to explore remains. To facilitate and prompt the next round of discussion, this final chapter describes some of the common and complementary themes about evaluation policy and outlines potential directions for next steps in understanding and influencing evaluation policy.

What Have We Learned?

This section focuses on three crosscutting issues: evaluation policy dimensions, evaluation policy instruments, and the political and economic environment for evaluation policy.

Evaluation Policy Dimensions. To guide study and development of evaluation policy, we need to know what dimensions or topics it encompasses. Trochim and Datta, in this issue, have started to define the limits and fill in the contours of the domain of evaluation policy. Trochim provides an eight-slice pie, while Datta draws on and adds to policy considerations found on the AEA Evaluation Policy Task Force (EPTF) Website (<http://www.eval.org/EPTF.asp>). When the two are compared, it appears that Datta's considerations begin to operationalize some of the slices in Trochim's pie. (Table 7.1 shows a rough correspondence between the dimensions in Trochim's framework and the issues identified by Datta.)

The absence of questions and issues that seem to line up directly with roles and use policies is not a comment on either the typology or the set of considerations. Instead, it is a two-pronged prompt for further exploration. We should ask whether these are the right categories, and at the same time we should think about how these categories can be further defined and articulated. Empirical research on what policies are already in place will be a critical step in refining our understanding of the domain of evaluation

Table 7.1. Match of Policy Types to Policy Questions and Issues

<i>Evaluation Policy Types</i>	<i>Policy Questions and Issues</i>
Evaluation goals	Evaluation definition (formal definitions and distinctions from functions such as program planning, monitoring, and performance measurement)
Evaluation participation (who is involved in <i>evaluation</i> ?)	Who is involved (who is involved in <i>evaluation policy development</i> ?)
Evaluation capacity building	Human resources (types of training, experience, and background required for evaluators) Resource distributions (distribution of capacity-building resources across diverse evaluation perspectives)
Evaluation management	Requirements of evaluation (when are evaluations required, what programs should have evaluations, how often are evaluations scheduled?) Evaluation budgets (standards for budgeting evaluation work)
Evaluation roles	
Evaluation process and methods	Evaluation methods (approaches recommended or required for what types of programs)
Evaluation use	
Evaluation of evaluation (meta-evaluation)	Evaluation ethics (policies for addressing ethical issues in evaluation)

policy. In the final section of this chapter, we identify specific ways to take these next steps.

Evaluation Policy Instruments. In his evaluation policy wheel, Trochim presents high-level policies as general statements that do not include operational specifications as the outer ring, with evaluation practice (the manifestation of policy) as the center. Moving from the outer ring to the center brings increasingly specific statements. At some point, the high-level policies are translated into policy instruments or strategies, which are in turn translated into practice. After reviewing several alternative classifications of policy instruments, Vedung (1998) proposes a set of three categories: “carrots” or economic incentives, “sticks” or regulatory constraints, and “sermons” or information.

In this volume, regulatory instruments are more frequently discussed than either economic or information strategies. For example, many of the details in the Government Performance and Results Act exemplify the stick, requiring annual collection and reporting of evaluative information on

agency performance. Another example is evaluation set-asides, such as the set-aside for evaluation in the Public Health Service. Set-aside legislation could be considered an economic incentive because it provides resources for evaluation. However, as a regulation that requires agencies to use funds for evaluation that they might prefer to put toward programs, it is more of a stick than a carrot. Evaluation steering groups or evaluation units would be considered organizational strategies rather than policy instruments in Vedung's analysis (1998). But, for example, the European Commission requirement to have evaluation steering groups can be considered a regulatory policy instrument. Preferences given to evaluation proposals using certain kinds of designs and support for evaluation capacity-building activities are examples of economic policy instruments.

The information instrument, or sermons, is restricted to "no more than pure transfer of knowledge, persuasive reasoning, or exhortations . . . offered to influence the public or some segment of the public to do what government deems desirable" (Vedung & van der Doelen, 1998, p. 103). Leeuw illustrates the use of information as a policy instrument in the government's exhortations about the need to increase evaluation. Information could also be an instrument for Chelimsky's recommendations for policies about the kinds of knowledge that evaluators should have (of agency history and culture). The policies might be translated by encouraging evaluators to understand their surroundings and furnishing the information needed to do so.

Regulatory, economic, and information instruments begin to articulate how evaluation policies move from general statements to practice. It is perhaps not surprising that examples of implementing evaluation policy echo in kind the previous literature on instruments of (substantive) policy. To influence evaluation policy, evaluators will need a sophisticated understanding of policy instruments.

The Political and Economic Environment. Evaluation policy is made and implemented in specific political and economic contexts, whether organizational, national, or global. Organizationally, the political environment can be a threat to the independence and credibility of evaluation units; thus evaluation policy is needed for protection. At a global level, the EU's political environment (a coupling of supranational and decentralized politics in a system of multiple nations with their own histories, cultures, and interests) has created a demand for evaluation for the purposes of transparency and accountability, which has in turn generated several policies related to use of evaluation and its linkage to decision making. The development of sound evaluation policy may be especially important in a context of the economic uncertainties (uncertainties that some might say are the result of an absence of evaluation policy, especially policies about evaluation goals and use). Leeuw gives an example of the role of the economic environment in evaluation policy in the Netherlands:

“reconsideration studies,” which were conducted primarily by internal ministry staff, were required to describe what would happen in terms of “goal achievement” if the budget for an intervention was cut by 20%. These examples indicate how evaluation policy reflects the political and economic environment and also has the potential to protect evaluation from political and economic threats.

Where Do We Go From Here?

Further conceptualization and description is needed to strengthen our ability to influence evaluation policy. This final section outlines next steps in conceptualization, description, and action.

Defining the Regions and Boundaries of Evaluation Policy.

Through evaluation policy, evaluators are influenced by and can influence an entity’s expectations for evaluation—its evaluation goals, its participants, its standards of quality, and so on. To understand the opportunities and limits of evaluation policy, we need to develop an understanding of the domain. An iterative process of conceptualization and description of the domain of evaluation policy could start with Trochim’s eight categories (see Table 7.1). Then, empirical investigations could address questions such as:

- Do these categories apply equally across evaluation contexts and types?
- Are any categories missing?
- Are the categories at the same level of generality?

To answer these and other questions, description and documentation of existing policies are needed. An evaluation policy archive could inform conceptualization of the domain of evaluation policy and constitute a basis for exploring questions about policy implementation.

Identifying Effective Evaluation Policies. Ideally, a taxonomy and documentation of evaluation policies establishes a foundation for action, specifically by helping evaluators identify, create, and influence better evaluation policies (and to discuss, debate, and define what “better” means). Our first task, then, is to get a handle on what kinds and levels of evaluation policies are most appropriate under what circumstances. For example, both the Netherlands and the EU seem to have some clear policies about evaluation goals, but decisions about evaluation methods are either made as informal policies or delegated to the practitioners (Leeuw, Stern, this issue). Is this the result of a conscious decision to give priority to goals policies by formalizing them and not others, or simply an evolutionary process in which policies of equal priority are formalized sequentially? Evaluation theorists and those who do research on evaluation could lead the way in deepening our understanding of effective

evaluation policy. In addition, evaluation journals could also encourage submissions on experiences with evaluation policies and the lessons drawn from those experiences.

As we collectively get a sense of the consequences of specific evaluation policies, our next task is to facilitate development of wise policies. Professional associations have a major role to play here. They can sponsor archives of evaluation policies, support professional development on evaluation policy at conferences and other events, and encourage networks for evaluators trying to influence policy. For example, AEA is about to release a Web-based platform that members interested in this kind of work could use to discuss and collaborate on evaluation policy research and development. In addition, professional associations can try to influence evaluation policy directly, as AEA has been doing for the last two years with its Evaluation Policy Task Force (EPTF). The EPTF is working with federal agencies on the executive branch side and with committee staffers and lobbyists in the legislative branch to try to influence federal policies on evaluation (see <http://www.eval.org/EPTF.asp> for more information).

As we raise consciousness about evaluation policy and encourage development of sound policies, we will also want to monitor the quality and implementation of the policies that are enacted. For this purpose, the kind of governmentwide audits that have been used in the Netherlands could furnish a model (Leeuw in this issue). A related task would be development of checklists and other measures to apply in these audits. Once the evaluation policy taxonomy has been further developed, it may amount to the basis for such measures.

Reprise

Evaluation policy shapes and defines our practice. As evaluators, we can influence the values expressed through evaluation policy, but only if we have a way to describe evaluation policies and policy instruments to policy makers, a sense of which evaluation policies are appropriate and effective, and facility with policy-influencing tactics. The ideas presented here and in the chapters of this issue of *New Directions for Evaluation* are only the first steps toward a more sophisticated understanding of evaluation policy and of strategies for developing and influencing it.

References

- Julnes, G., & Rog, D. J. (Eds.). (2007). Informing federal policies on evaluation methodology: Building the evidence base for method choice in government sponsored evaluation. *New Directions for Evaluation*, no. 121. San Francisco: Jossey-Bass.
- Vedung, E. (1998). Policy instruments: Typologies and theories. In M.-L. Bemelmans-Videc, R. C. Rist, & E. Vedung (Eds.), *Carrots, sticks, and sermons: Policy instruments and their evaluation* (pp. 21–58). New Brunswick, NJ: Transaction.

Vedung, E., & van der Doelen, F. C. J. (1998). The sermon: Information programs in the public policy process—Choice, effects, and evaluation. In M.-L. Bemelmans-Videc, R. C. Rist, & E. Vedung (Eds.), *Carrots, sticks, and sermons: Policy instruments and their evaluation* (pp. 103–128). New Brunswick, NJ: Transaction.

LESLIE J. COOKSY, incoming president of the American Evaluation Association, is an associate professor at the University of Delaware where she directs a graduate program in evaluation.

MELVIN M. MARK is professor and head of psychology at Pennsylvania State University. He has served as president of the American Evaluation Association and as editor of the American Journal of Evaluation (now editor emeritus).

WILLIAM M. K. TROCHIM is professor of policy analysis and management at Cornell University and is the director of evaluation for the Weill Cornell Clinical and Translational Science Center, the director of evaluation for extension and outreach, and the director of the Cornell Office for Research on Evaluation.