

Independent Evaluation and the Tension Between Accountability and Learning

What should be the main goal of evaluation? Is it to hold programs and policymakers accountable for accomplishing their intended goals? Is it to help program managers and policymakers learn how to do their work better, namely to minimize errors and seek constant improvement? As the title of this chapter suggests, there can be a persistent tension about the relation of accountability and learning in evaluation studies. The debate between these two distinct objectives continues unresolved in the literature, with the traditional framework in evaluation putting it as an “either-or” proposition—you can focus on accountability or on learning, but you cannot have both.

This chapter lays the groundwork for highlighting that the “either-or” dichotomy is not necessarily valid. It focuses on the accountability and learning dimensions of evaluation and poses the question of whether the utility of evaluation needs necessarily to be framed in terms of one or the other. The discussion sets the stage for the positing in the next chapter that both purposes can be served within a single study. To begin, in this chapter, we focus on the notions of accountability, organizational learning, and their relation both to each other and to evaluation.

Before proceeding, we briefly address here two other motivations for evaluation identified by Vedung (1998). The first of these is to politically or symbolically legitimate the functions of the organization. Here the use is to show citizens and stakeholders that the functions and activities of the organization are taken seriously and hence are systematically evaluated by professionals. The second is to permit the postponement of decisions. Policymakers who are under pressure to decide on an issue can claim that it is too soon to do so because an evaluation is under way. Such a postponement can work to the benefit of the policymaker by letting time pass, hopefully letting the decision become less salient and perhaps move to the desk of someone else.

On the first of Vedung’s categories, the use of evaluation to legitimate the function of an organization, one of IEO’s purposes as established in

its terms of reference is precisely to “strengthen the Fund’s external credibility” (IMF, 2000b). This objective remains to this day as part of a triad, along with accountability and learning, as the key components of IEO’s mandate.¹ The Ocampo Report acknowledged that “there was a strong consensus, from inside the IMF, from national governments, and from external stakeholders, that the IEO had strengthened the IMF’s external credibility” (Ocampo, Pickford, and Rustomjee, 2013: 3). Furthermore, the panel concluded that strengthening the Fund’s external credibility had largely been achieved as the result of the exercise of the other two mandates: namely, enhancing the learning culture of the Fund and supporting its institutional governance and oversight.

With regard to the use of evaluation for postponing decisions, we would venture that the IMF has never used the conclusions or recommendations of an IEO evaluation as a justification for stalling its agenda or work. We would also conjecture that if ever the Fund has postponed taking action, it would have done so only in anticipation of an upcoming IEO evaluation report, in order to see whether IEO’s analysis would provide additional insight. It has been IEO’s experience that IMF management and staff have occasionally tried to preempt an IEO evaluation by taking action on a topic under evaluation (described in the 2006 Lissakers Report as “front-running”). This strategy, that aims to address some of the possible findings of an IEO report before the evaluation is completed, could be seen as a positive outcome of IEO’s presence but may also mute the eventual findings of the IEO report, which may yield different or more critical conclusions.

Accountability and Learning

To many within the evaluation discipline, the primary purpose of evaluation is to help ensure accountability. Other practitioners consider that evaluation also has the responsibility to highlight the value and importance of organizational learning. For some, the prospects of reconciling the two is problematic (Lehtonen, 2005) while others do not see so much complementarity as they see these two objectives operating in two different domains.

Bemelmans-Videc, Perrin, and Lonsdale (2007: 250) argue that “[t]raditional forms of accountability are often viewed as less concerned

¹The IEO’s original Terms of Reference (IMF, 2000b) also included as a mandate of the Office to “promote greater understanding of the work of the Fund throughout the membership.” However, following the recommendation of the second external evaluation of the IEO, this objective was dropped, given that the panel had determined that this had been achieved.

with learning than with punishment.” Others have noted that “a primary focus on accountability brings with it a strong focus on rigor, independence, replicability, and efficiency, whereas a focus on learning emphasizes stakeholder ‘buy-in’ and an evaluation process which leaves space for discussion and lesson-drawing” (Lonsdale and Bechberger, 2011: 268). Stated somewhat differently by the OECD (2001: 68), “[t]hese two objectives are not necessarily incompatible . . . but they are sufficiently different to merit separate consideration.”

There is widespread discussion in the literature on how evaluative work may contribute to learning in an organization. Howlett and Ramesh (1995), for example, see policy evaluation as part of a process of learning in which policies evolve mainly because of the recognition of past successes and failures and deliberate efforts to emulate successes and prevent failure. Learning from success is increasingly understood as a powerful means of learning (cf. Nielson, Turksema, and van der Knaap, 2015).

Learning from evaluation is, of course, far from inevitable or straightforward. It may well depend on a range of factors including organizational capacity, the cultural value of learning in the organization, the approaches used, the authority of those carrying out evaluations, the authority of those receiving the evaluations, the appropriateness of timing, luck, and whether there are forces working against learning. Some authors also emphasize the importance of ongoing links between evaluators and those whose work is being evaluated, the incremental and iterative nature of learning, and the value of learning from past evaluations (Preskill and Torres, 2000).

Consider the following question by Picciotto:

Can an internal independent evaluation function be designed to promote organizational accountability—or is it condemned to be an empty ritual? Is it conducive to organizational learning—or does it produce a chilling effect that inhibits adaptation to changing circumstances? (Picciotto, 2013: 18).

Independent evaluation units were created to help protect the credibility and legitimacy of the management process. Evaluation, when properly done, enhances the credibility of an organization’s management when policies, processes, and programs are evaluated. But if adequate care is not taken in the design of an evaluation unit, its efforts can be sidelined by the rest of the organization, and thus the benefits of evaluation, especially as a learning device, are nullified. Hence it is important to be aware that the mere establishment of an independent evaluation function within an organization creates barriers that work

in unison to resist the permeation and incursion of the evaluation unit into the rest of the organization (Mayne, 2008).

Even with the unavoidable creation of such barriers, the evaluation unit must strive to promote accountability and learning, and in a constructive and positive way, while maintaining its independence. By doing this, it can somewhat ease the inherent tension between accountability and learning and the organization will more easily reap the benefits of independent evaluation. To support this argument, Picciotto adds:

... deeply adversarial attitudes and “name and shame” approaches rupture contacts with decision makers, restrict access to tacit knowledge, inhibit professional exchanges and increase resistance to adoption of evaluation recommendations. They lead to isolation, a lack of intellectual leverage, and a chilling effect on organizational learning. This is why diminishing returns set in when evaluation independence assumes extreme and antagonistic forms (Picciotto, 2013: 22).

Consider also this statement by van der Meer and Edelenbos:

... there is an increasing emphasis on transparency, measurable results, and accountability. Policy documents should specify clear goals, the attainment of which should be measured by unequivocal (and if possible quantitative) indicators. Policy-makers should be held accountable for the results thus assessed. Evaluation, therefore, should assess efficiency, output, and outcomes of policies against their (initial) goals (van der Meer and Edelenbos, 2006: 202).

The implications of this statement need to be made explicit: (i) the purpose of programs and policies needs to be clear; (ii) intended objectives need to be measured as precisely as possible; (iii) those who are responsible for achieving the objectives should be held to account; and (iv) accountability should not aim at “shame and blame” but rather focus on an assessment of results, particularly outcomes. An additional dimension of accountability not mentioned in this quote also merits attention: that of what is learned and acted upon from evaluation studies. Accountability studies can provide an array of analyses to help managers do their work better. As Mayne notes:

Finding out why things are or are not working and seeking ways to improve programs and policies is what most evaluations are all about. This aligns well with the learning aspect of accountability. Managers want to learn how to improve their programs and policies and should be eager to demonstrate they have done so (Mayne, 2007: 79).

Thus, while accountability provides an opportunity to appraise whether and how an activity is being done and what the consequences are of doing it well (or not), it also provides the opportunity to learn how to do better.

In Mayne's own words (2007: 81): "Accountability focused more on learning than blaming provides greater potential for evaluation to play a meaningful role."

Thus, evaluation can achieve a position in which it provides both accountability and learning. In this endeavor, Lehtonen emphasizes the need to clearly define the roles of each:

There is an obvious tension between the two perspectives on the use of evaluation and they are often seen as irreconcilable. However, most authors seem to recognize that both providing accountability and enhancing learning are essential elements in the endeavor to promote 'social betterment' through evaluations. . . . The challenge is therefore not to choose between the two, but to look for complementarity through clearly defining the roles of the two approaches (Lehtonen, 2005: 170–71).

Even if evaluation studies provide both elements of accountability and learning, how can an organization overcome its resistance to change? Inertia, and entrenched interest in maintaining the status quo, are powerful forces to overcome (Perrin, 2015).

As will be discussed in [Chapter 9](#), the leadership of an organization has an important role to play in creating the conditions under which evaluation studies affect learning—that is, in overcoming the organization's inherent resistance to learn and thus change. As Perrin notes:

Leadership from the top is needed to bring about and to support needed organizational renewal and change. There is still limited, but an increasing array of resources available about how to manage for outcomes in a way that embraces complexity. Organizations that remain static and fail to evolve and improve quickly become out of date and may struggle to survive, at least in the long run (Perrin, 2015: 14).

New Wine in New Bottles—Or a Better Approach to Accountability

Perrin, Bemelmans-Videc, and Lonsdale (2007) make the point that a new and different way of thinking about accountability can be developed, especially when organizations are striving to achieve outcomes in the context of a complex policy environment and are facing a variety of complicating factors. As summarized by Perrin (2015), three essential characteristics define this different approach to accountability:

- (i) A primary orientation towards results rather than on process;
- (ii) A focus on continuous and responsive learning; and
- (iii) A dynamic rather than a static approach that reflects the complexities and uncertainties inherent in most public policy areas.

Perrin notes:

This model of accountability involves holding programs accountable for asking the tough questions about what works and why, innovating and engaging in risk taking rather than playing it safe, and for seeking—and using—feedback. Holding programs accountable for asking the difficult questions, doing and using evaluations, and demonstrating use of learning—such as through changes in policies and in program approaches, may represent a harder standard than demonstrating compliance with procedures as with traditional accountability. In short, programs should be accountable for demonstrating good management and for keeping in view outcomes, which includes (but definitely is not limited to) a true results orientation (Perrin, 2015: 15).

Perrin, Bemelmans-Videc, and Lonsdale (2007) essentially argue for a new framework governing how one approaches the notion of accountability. What will be required is to transform the traditional compliance-oriented accountability approach into one that is nimble, learns from mistakes, follows up with corrective action, stops trying to establish and lay fault, and gives up on the “shame and blame” approach. This amounts to a transformation into a culture of learning. If reforms are not undertaken, if little to no emphasis is placed on moving towards that culture of learning, and if accountability is more and more understood as rhetorical, the organization’s performance will not improve and the confidence of those relying on evaluation will decline.