

# “Show me the Numbers”: Examining the Dynamics Between Evaluation and Government Performance in Developing Countries

LESLI HOEY\*

*University of Michigan, United States*

**Summary.** — This paper examines the dynamics between monitoring and evaluation (M&E) and government performance in developing countries, where M&E systems are expanding rapidly. Findings in Bolivia suggest that approaches to M&E can lower staff morale, create burdensome paperwork, blind managers to operational problems and emerging innovations, and reinforce self-censorship, contributing to the very problem M&E is intended to solve. Crafted appropriately, M&E can instead become a tool to build practical judgment, increase staff motivation, and improve implementation incrementally. Ultimately, these findings contribute to efforts to design M&E that can support staff working under complex working conditions.

© 2014 Elsevier Ltd. All rights reserved.

*Key words* — monitoring and evaluation, policy implementation, development planning, malnutrition, health policy, Bolivia

## 1. INTRODUCTION

This paper examines the dynamics between monitoring and evaluation (M&E) and government performance in developing countries, where M&E systems are expanding rapidly (Ernesto, Shand, Mackay, Rojas, & Saavedra, 2006; EvalPartners, 2014). Rather than asking about the technical quality and rigor of M&E being used (Fukuda-Parr, Greenstein, & Stewart, 2013), my questions build on research about the conditions under which such systems improve or erode development policy implementation (Hood, 2012). Based on a case study of Bolivia’s Zero Malnutrition (ZM) program, I suggest that mid-level managers may cling to collecting information about externally defined, quantitative indicators of staff performance as a *reaction* to complex social change processes—as coping mechanisms that give the illusion of controlling implementation. In these situations, evaluation can obscure operational issues, create burdensome paperwork, blind managers to emerging innovations, and reinforce self-censorship, contributing to the very problem M&E is intended to solve. On the other hand, where managers use M&E in ways that help build practical judgment about how to improve implementation, they create an environment conducive to learning—building motivation and trust—and engage more diverse actors and types of knowledge in analyzing problems and negotiating solutions. Ultimately, these findings contribute to literature aiming to reconsider how to design M&E to support staff working under complex conditions (Rogers & Fraser, 2014).

Efforts to institutionalize government-based monitoring and evaluation<sup>1</sup> (M&E) systems in developing countries have grown considerably over the past decade in response to the Millennium Development Goals (Sayedoff, Levine, & Birdsall, 2006), Poverty Reduction Strategy Papers (Holvoet, Gildemyn, & Inberg, 2012), and the Paris and Accra Declarations (High Level Forum, 2008; OECD/DAC Organization for Economic Coordination and Development/Development Assistance Committee, 2005). Each of these initiatives calls for more “country-owned” development (Hyden, 2008) as well as monitoring and evaluation of donor investments and public policies (Thomas, 2010). In response, international institutions have launched numerous evaluation

networks (IOEC, 2014) and initiatives to build development evaluation capacity (Mackay, 2006; Naidoo, 2013; Sayedoff *et al.*, 2006). There is increasing evidence of “country-led—rather than donor-driven—efforts to institutionalize M&E” (May, Shand, Mackay, Rojas, & Saavedra, 2006, p. xi; Imas & Rist, 2009), and the demand for evaluators is growing; as of 2012, there were 138 national associations of professional evaluators representing 110 countries (EvalPartners, 2014), up from only five in 1990 when associations existed only in North America, Europe, and Australia (Donald, 2006).

At the heart of this exponential growth is a belief that M&E serves a variety of purposes: to hold actors accountable, identify policy options proven to work, and to improve the effectiveness of interventions during implementation (Hood, 2012; IOEC, 2014). While researchers—most notably those with the Abdul Latif Jameel Poverty Action Lab (J-PAL) (Kremer and Glennerster, 2012)—are showing that rigorous (e.g., randomized control trial) evaluations can help identify effective international development strategies, barriers still exist to support wider adoption of evidence-based policy during the planning process (Dhaliwal and Tulloch, 2011). Moreover, there is no standardized approach for building country-level capacity to mainstream M&E during the implementation phase (Goldberg and Bryant, 2012). One of the major debates is about whether certain evaluation models may contribute to the very problem they are intended to solve

\* This study was supported primarily through an Inter-American Foundation Grassroots Development Fellowship as well as Field Research Grants awarded through Cornell University’s Tinker Program in Latin American Studies, Einaudi Center for International Studies, and International Studies in Planning. The study’s sponsors played no role in the design of the study, in the collection, analysis, and interpretation of the data, in the writing of the report, or in the decision to submit the paper for publication. I am grateful for the support, assistance, and participation of staff at national, departmental, and municipal levels of Bolivia’s Ministry of Health and Zero Malnutrition Program and stakeholders from Bolivia’s international nutrition policy community. I also thank John Forester, David Pelletier, and Mark Constanas for their feedback on an early version of this report, as well as two anonymous referees for their valuable comments and suggestions. Final revision accepted: December 18, 2014.

(Hood, 2012; Rogers and Fraser, 2014; Westley, Zimmerman, & Patton, 2007). Non-governmental organization (NGO) staff and other development scholars have long argued that dominant forms of M&E systems focused on linear program designs, pre-determined, quantifiable indicators, and efficiency outcomes can discourage adaptation and innovation and encourage short-term and risk-averse projects (Chambers, 2010; Eyben, 2010; Natsios, 2010; Patton, 2010). Esser (2014) also raises concerns that the “country-ownership” and “aid harmonization” discourse supported by the Accra and Paris Declarations is moving the locus of accountability onto countries, so that they must take the responsibility for failures and successes, even as donors maintain their negotiating power over the types of interventions and M&E systems that countries must agree to in order to receive funding, what Esser calls a form of “expost-conditionality” (p. 51). The rapid rise in so-called “country-led” M&E systems, in this light, suggests that much of the trend is focused on pleasing donors (Eyben, 2010; Sjöstedt, 2013), rather than on improving government performance.

Regardless of the reasons, the growing evaluation agenda in developing countries suggests that a clearer understanding is needed about the ways in which M&E approaches sometimes complicate, rather than improve, complex interventions, even as we identify more appropriate M&E strategies for the types of problems development actors face. In what follows, I explain how different forms of M&E should be expected to affect policy implementation under different conditions, based on theories about complexity, bureaucracy, organizational change, and behavior economics. The remaining sections outline why Bolivia’s ZM program offers a useful lens through which to explore these dynamics, the methods used in this study, and finally, how the findings offer lessons for re-designing development M&E systems.

## 2. THEORIES ABOUT M&E EFFECTS ON POLICY IMPLEMENTATION

At the heart of the many debates about the supposed benefits and drawbacks of evaluation are different understandings of the policy process that development actors face, with divergent implications for the form of M&E that should be appropriate. M&E proponents who believe that evaluation can inspire order and improve development results are often drawing on Max Weber’s view of bureaucracy as a formal-rational system. Weber believed it was the normative appeal of rational-legal authority—the impersonal laws, procedures, and rules—that would compel employees to perform, because it offered a depoliticized, fair, stable, and predictable way of ordering society, rather than decision-making based on subjective beliefs, values, tradition, faith or “charismatic gifted persons” that was more common at the time (Gerth and Wright Mills, 1970, p. 199). Measurement systems play a key role in supporting this form of objective decision-making, allowing managers to track each worker to ensure results, as Weber wrote: “the performance of each individual worker is mathematically measured, each man becomes a little cog in the machine” (Weber in Mayer, 1956, p. 127) making “possible a particularly high degree of calculability of results for the heads of the organization” (Weber, 1978, p. 223).

The Weberian model persists, and the form of disciplining M&E that goes along with it, because it works in many situations (Stacey, 1996). This view of bureaucracy should produce expected results, complexity theorists argue, when public problems are “simple”—like baking a cake—when there are clear and agreed solutions to the problem, and the solution

can be perfected through repetition and strict adherence to a recipe (Westley *et al.*, 2007). Complexity theorists and organizational change scholars also agree that a Weberian bureaucracy can function well when problems are “complicated”—like sending a rocket to the moon—where confounding factors can be reduced with enough information and coordination through centralized decision-making (top-down bureaucracies that rely on expertise) and technical rationality (based on planning, evaluation targets that allow for quickly identifying and “fixing” weaknesses in inputs, and sanctions and incentives to “command and control” staff into a coherent system) (Elmore, 1980, p. 605; Glouberman and Zimmerman, 2002; Mazmanian and Sabatier, 1983, p. 20; Westley *et al.*, 2007).

Scholars argue, however, that problems arise when Weberian bureaucracy—with rigid rules and M&E used to control behavior—is used in a situation where problems are “complex” (Guijt, 2007; Westley *et al.*, 2007) or “wicked” (Rittel and Webber, 1973). These include problems where causes are multidimensional and dynamic, and a multiplicity of stakeholders have conflicting perspectives about solutions. Hodson, Martin, Lopez, and Roscigno (2012) contend that this can turn institutions into “Kafkaesk” bureaucracies, where the norm is “divergent goals, unwritten rules, patrimonialism” and “chronic states of contradiction and confusion” (Hodson *et al.*, 2012, p. 265). Their analysis of 160 institutional ethnographies showed that rule breaking occurred routinely in 60% of organizations, while 86% showed widespread evidence of at least one of the “Kafkaesk” bureaucracy characteristic, leaving only 14% operating entirely through Weber’s formal-rational model (p. 265). The authors conclude that “mock bureaucracies”, full of “confusion, deceit, conflict and personal power” (p. 257), or situations where management and staff informally agree to break the rules, should be expected as the norm, rather than the exception to the Weberian rule (p. 260). This is akin to the type of rule-breaking behavior Lipsky (1980) found among street-level bureaucrats, or Friedmann (1993) arguments that implementation is inevitably a political act, displacing some existing practice, resource, time, staff, decision-making power, and more.

Behavior economics and organizational change research reinforce these arguments, showing how incentives and fines can cause staff to do the opposite of what the supervisor intended under certain conditions (Bowles & Polania-Reyes, 2012; Osterloh, Bruno, & Homberg, 2007). This may be particularly true among civil servants who have been found to “have a greater interest in altruistic activities and socially desirable outcomes” (Osterloh *et al.*, 2007, p. 11). Such intrinsic motivation has been shown to foster creativity, speed learning, improve conceptual understanding of the problem and solution, and encourage a more holistic approach (Hodson *et al.*, 2012; Osterloh *et al.*, 2007). However, introducing rewards or sanctions—including negative or positive feedback from M&E—can lead a person to lose their intrinsic motivation and interest in the immediate goal (e.g., to deliver a service) and shift their “locus of causality” externally to do the activity on the basis of the reward or punishment, so that “you get what you measure” (Bowles and Polania-Reyes, 2012; Hodson *et al.*, 2012; Osterloh *et al.*, 2007, p. 6). Especially when evaluations are used to critique and punish, this can lead staff to blame negative evaluations on outside factors or defend their actions rather than learn from them (Argyris and Schön, 1996; Frey, 2010, p. 17). The meaning an employee attributes to the incentive also matters, so that if they view it as a form of control or if they believe the supervisor distrusts them, they may intentionally perform worse to exert their sense of autonomy (Bowles and Polania-Reyes, 2012).

Download English Version:

<https://daneshyari.com/en/article/7393785>

Download Persian Version:

<https://daneshyari.com/article/7393785>

[Daneshyari.com](https://daneshyari.com)