CHAPTER 5

The Logic of Evaluation Independence and Its Relevance to International Financial Institutions

ROBERT PICCIOTTO

> *There is much to be done in the design of institutions to reconcile the values of authority and responsibility.*
>
> —Kenneth J. Arrow

This chapter is about the role of independent evaluation in organizations, in particular in international organizations. Evaluation must be free from external pressure if it is to produce meaningful evidence in support of institutional learning and effective and accountable decision making. Thus, a critical aspect of independent evaluation is that evaluators do not report to those in charge of the programs or practices that are being evaluated. But how does independent evaluation strengthen organizational effectiveness and facilitate organizational reform? Why is it needed to generate objective scrutiny of policies, programs, and processes? Why is attesting to the validity of self-evaluation processes important to achieve accountability and learning in public organizations?

These questions are especially relevant for international financial institutions (IFIs).[1] To learn from experience and keep up with the rapid changes taking place in their operating environment, IFIs need objective evidence regarding the effectiveness of their operations. Further, they need to nurture public trust in the integrity of their decision making. An independent evaluation function can contribute to these goals by providing a credible assessment of their performance.

The chapter is organized as follows. The first section makes the case for evaluation independence and outlines its essential features. The second section shows that independent evaluation is a powerful instrument of organizational learning. The third section identifies the complementary roles of independent and self-evaluation. The fourth section addresses the size of the evaluation function. The

[1]IFIs are public sector organizations involved in financial operations, and their shareholders are national governments. The largest and most influential are the IMF and the World Bank. At IFIs, independence for evaluators requires that they not report to the head of the organization and not be part of its management structure. Thus, the head of the evaluation unit is typically accountable to the members/owners via some form of loose reporting to the organization's board. This chapter also discusses other aspects of independence, for example, organizational and behavioral independence, and safeguard mechanisms.

fifth section explains why independent evaluation is a critical ingredient of good governance in international financial institutions. The final section concludes.

Evaluation Independence: Why and How?

Independence adds credibility to evaluation processes and judgments, and contributes to organizational effectiveness. Organizations are created to resolve conflicts between individual and collective goals, manage information flows, coordinate actions through a nexus of contracts and keep transaction costs in check. To enforce the codes of conduct that make this possible requires authority. But authority is only tolerated if it is legitimate—that is, if those entrusted with running the organization are perceived as responsible stewards of the resources entrusted to their care. In turn, to hold authority responsible requires mechanisms that help to ascertain whether errors in decision making were due to circumstances over which the organization had no control or whether risks could have been managed differently so that these errors would have been avoided (Arrow, 1974).

This is where independence in evaluation comes in. It protects the integrity of the assessment process, enhances its credibility, minimizes bias, and provides fresh perspectives on the policies and programs being evaluated (Mayne, 2008). It informs the adaptation of internal processes and protocols through organizational learning. To these ends, effective evaluation processes are designed and managed so as to assess the worth of organizational activities in a fair, valid, and accurate fashion.

Such goals cannot be met if evaluation takes place within an authorizing environment that:

- constrains information so that evaluation products cannot have any critical content;
- controls the content of the evaluation program so that it does not contribute new knowledge;
- delays the evaluation process (or the disclosure of evaluation results) until after the decisions that might have been informed by the evaluation have been made; and
- induces evaluators to focus on insignificant aspects of the program or policy being evaluated.

Characteristics of an Independent Evaluation Function and Evaluators' Competencies

Evaluation quality without independence lacks credibility. Conversely, independence on its own does not guarantee evaluation quality: relevant skills, sound methods, adequate resources, and transparency are also required.

Good evaluation calls for practitioners who are curious, skeptical, and hungry for evidence. But evaluation independence also implies distinctive personal characteristics, attitudes, and behaviors that reach beyond expert knowledge and experience and involve such qualities as loyalty, perseverance, and courage. These dispositions are reflected in a fair, balanced, and self-confident approach that does not shrink from pointing out problems and performance shortfalls but also recognizes success and achievement.

The *Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Service Pronouncements* by the International Federation of Accountants (2010) distinguishes between independence of mind ("the state of mind that permits the provision of an opinion without being affected by influences that compromise professional judgment, allowing an individual to act with integrity, and exercise objectivity and professional skepticism") and independence in appearance ("the avoidance of facts and circumstances that are so significant that a reasonable and informed third party, having knowledge of all relevant information, including safeguards applied, would reasonably conclude [that] integrity, objectivity or professional skepticism had been compromised").

Similarly the *Yellow Book* of the General Accountability Office of the United States sets out criteria for auditing and evaluation that include "an independent attitude and appearance" as a desirable characteristic. However, evaluators should not be so detached as to shirk interaction with program managers, staff, or beneficiaries.

Institutional Safeguards Are Needed to Protect Evaluators' Independence

Independent evaluators cannot be expected to survive long in an organizational context that puts their careers, their reputations, or their livelihoods at risk. To give them protection to carry out their demanding and stressful mandate calls for special institutional safeguards to protect the evaluation function. Evaluators need to be shielded from external threats to their impartiality. They must be given full access to the information they need to carry out their work. And immunity from capture by any of the parties that share program management responsibility is fundamental.

Evaluation independence is a "hit and miss" affair unless the organization that contracts the evaluation provides safeguards that guarantee this independence. If evaluators are commissioned to carry out an evaluation, the process is governed by a contract and the evaluators' work is guided by voluntary guidelines issued by evaluation associations. In such cases, the degree of independence largely hinges on whether the evaluation commissioner is free of allegiances and devoid of interests in the policy or program being evaluated and on whether he/she is genuinely trying to find out whether a policy or a program works.

Safeguards for evaluation independence have been codified for cases where the evaluation is carried out or contracted by a unit embedded within an organization. The Evaluation Cooperation Group[2] has identified four major dimensions of evaluation independence (reproduced in the annex to this chapter):

---

[2]The Evaluation Cooperation Group is composed of evaluation heads among the IFIs. It is tasked with the harmonization of evaluation methods and the codification of good evaluation practices. See https://wpqr1.adb.org/LotusQuickr/ecg/Main.nsf/h_9BD8546FB7A652C948257731002A062B/AA95B62CF943F82E4825774B003A82E0/?OpenDocument&Form=h_PageUI.

- *Organizational independence* ensures that evaluation staff are not controlled or influenced by decision makers who have responsibility for the activities being evaluated and that, within ethical and legal constraints, they have full access to the information they need to fulfill their mandate.
- *Behavioral independence* measures the extent to which the evaluation unit is able and willing to set its work program, produce high-quality and uncompromising reports, and disclose its findings to the board without management-imposed restrictions.
- *Conflict-of-interest safeguards* guarantee that current, immediate future, or prior professional and personal relationships and considerations are not allowed to influence evaluators' judgments or create the appearance of a lack of objectivity.
- *Protection from outside interference* keeps the evaluation function free to set its priorities, design its processes and products, reach its judgments, and administer its human and budget resources without intrusion by management.

These evaluation independence criteria are interrelated. Protection from outside interference is the object of organizational independence. Conflicts of interest are frequent, without organizational independence. Behavioral independence is a function of organizational independence as well as of avoidance of conflicts of interest and protection from external interference.[3]

Being External Does Not Guarantee Independence

External evaluation is often equated with evaluation independence. Yet external evaluators suffer from inadequate understanding of the operating context. Furthermore, their judgment may be impaired or threatened if their services are retained by the managers in charge of the activities that are being evaluated: fee dependence is a major threat to the integrity of the evaluation process.

By contrast, internal evaluations funded and controlled by the organization's supreme governance authority are protected from management influence while enjoying proximity to the programs being evaluated. Such evaluations are more likely to overcome "information asymmetries" while protecting the objectivity of the evaluative process.

---

[3]The *Glossary of Key Terms in Evaluation and Results-Based Management* issued by the Development Assistance Committee of the Organization for Economic Cooperation and Development specifies that an evaluation is independent when it is "carried out by entities and persons free of the control of those responsible for the design and implementation of the development intervention." It also indicates that independent evaluation presumes "freedom from political influence and organizational pressure," "full access to information," and "full autonomy in carrying out investigations and reporting findings."

Independence Is Not Isolation

Independence and quality depend on the extent to which evaluations are:

- *Critical*—able and willing to judge performance objectively and transparently;
- *additional*—making a distinctive contribution to operational knowledge creation or dissemination;
- *timely*—delivering operations evaluation findings and lessons early enough to inform decision making; and
- *material*—focusing deliberately on topics and issues that have substantial relevance to operational effectiveness.

Some evaluators argue cogently that having no connection or shared experience with the intended users of evaluations constrains evaluators' access to information, evokes resistance, and inhibits learning.[4] Their concern underlines the challenge of combining evaluation rigor and objectivity with fairness and sensitivity. Good evaluators are able to combine intellectual detachment with empathy and understanding. They are able to engage with diverse stakeholders and to secure trust while maintaining the integrity of the evaluation process.

Contestability of independent evaluation findings and principled, respectful disagreements about recommendations make for a healthy evaluation culture. Conversely, 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 the adoption of evaluation recommendations. They lead to isolation and a chilling effect on organizational learning.

The Benefits Generated by Independent Evaluation Are a Function of the Organizational Context

These benefits are especially large in open and accountable working environments where innovation and creativity are rewarded. In such environments, evaluation independence evokes public trust, protects the learning process, and induces program managers and stakeholders to focus on results. Along with quality assurance and ethical guidelines, evaluation independence is an essential feature of organizational credibility.

Evaluation and the Public Sector

In the private sector, stakeholders can use market mechanisms to express their dissatisfaction with poor performance. Disgruntled shareholders can sell their shares and unhappy consumers can shift from one brand or supplier to another. This generates market signals that induce private sector managers to modify their behavior. In

---

[4]Jane Davidson, an eminent New Zealand evaluator, puts it this way: "If we don't know you and trust you, if you don't have any connection with us, then why would we share insights with you?"

the economic marketplace, private suppliers serve fairly homogeneous needs that are readily translated into measures of private merit and worth. And the market provides private corporations with ready access to information about their performance.

In the public sector, reliable performance measures are much harder to secure. Public policies and programs meet varied interests whose satisfaction is not easily translated into indicators. Good evaluations take account of the concerns of diverse stakeholders and strike appropriate trade-offs among them. The main recourse of disgruntled stakeholders is the "voice" option. The effectiveness of this option hinges on ready access to timely, valid, and reliable markers of performance. This is where evaluation comes into its own. Without it, the feedback mechanism of the political marketplace is intermittent—from one electoral cycle to the next.

Furthermore, public action impacts and works through the private sector and the civil society. Policymakers and managers of social programs cannot create value through exclusive reliance on factors within their control. Their performance hinges on their capacity to respond nimbly and appropriately to changes in the external operating environment and to shape the perceptions of the politicians and decision makers who control their fate (the authorizing environment).

Evaluation Connects Program Knowledge with Policymaking and Organizational Strategy

It follows that the external operating environment should help evaluators generate the indicators needed to inform strategic and policy decisions. Accordingly, the evaluation function should design and use processes that connect it to the beneficiaries of public sector interventions. Feedback from these constituencies can be secured through a variety of participatory evaluation tools (such as focus groups and surveys) and increasingly through social media technologies. The signals thus received need to be interpreted fairly and judiciously before they are transmitted in a timely and user-friendly fashion to the deciders.

In Order to Contribute to Public Policies and Organizational Strategies, Evaluation Must Be Independent

Evaluation cannot fulfill its potential unless it connects effectively to corporate management, the supreme authorities that govern the organization, and the broader society. But in maintaining these connections it should maintain its objectivity, exercise full freedom of inquiry, and resist capture. This is imperative since it cannot deliver on its mandate if it is perceived to be subservient to program managers or to a particular political constituency or group.

Independent Evaluation Enhances Organizational Accountability

All managers, whether in the private or public sector, are accountable for results to the bodies that control their funding. In particular, they need to show that they are responsible stewards of the resources entrusted to them. In the private sector, revenues derive from the sale of goods and services to individual consumers—or from private capital markets that judge the prospective value of such sales. But in the public sector, whose original source of funds is the taxpayer, it is up to civil servants and ultimately to politicians to make collective choices about how public resources are allocated. Hence the key to accessing resources in the public sector is a valid and authoritative narrative regarding the creation of public value that is judiciously targeted to persons in authority.

Ready measures of public sector performance have well-known disadvantages. The difficulty of measuring social value explains why simple and inexpensive output measures and budget coefficients, rather than indicators of outcomes and impacts, have traditionally dominated public sector management. The ready measures have major and well-known drawbacks. They do not measure results and they can easily be manipulated. Hence, the information provided by public sector managers about the effectiveness of their work needs independent validation. This is where independent evaluation comes in. It is to a public sector organization what the auditing of corporate accounts is to a private firm.

This is why there is in practice little disagreement about the close link between independent evaluation and organizational accountability. If evaluation fails to enhance accountability it is because it lacks validity due to its poor quality and/or because it fails to meet the criteria that make evaluation truly independent. Far less obvious is the role of independent evaluation in enhancing organizational learning, discussed next.

The Role of Evaluation in Promoting Organizational Learning

To adapt successfully to evolving circumstances, learning organizations secure the right competencies and nurture the right skills through relevant individual learning by their staff, and they resolve the core dilemmas of collective action through controls and structures that leave opportunities open for new ideas. Evaluation fosters the adaptability of these organizations by focusing the attention of management on results—and hence on the appropriateness (or obsolescence) of the goals, policies, or protocols that underlie current organizational behavior.

Organizational Learning: Dilemmas of Collective Action

Learning challenges current ideas and preconceptions. It leads to new insights and encourages the acquisition of new concepts and mental models that serve as useful guides to action. Cognitive science, educational psychology, and practical experience show that learning often requires external intervention. It may be triggered by unexpected events or external threats. It may also be nurtured through interaction with experienced, knowledgeable, external agents through coaching, training, or formal education.

Organizational learning is not the same as individual learning. To be sure, it cannot be divorced from individual learning since individuals manage and operate organizations. But organizations are mostly driven by collective actions shaped by hierarchy, protocols, and precedent. Beyond knowledge acquisition, organizational learning requires attitudinal shifts and behavioral changes. Not all changes in behavior result from knowledge acquisition or experience; some result from conditioning through fear and reward, respect for authority, or a desire to conform. Accordingly decision makers within organizations are conditioned to comply with "rules of the game" shaped by legal, procedural, and traditional constraints and by time-honored customs that favor habitual ways of doing things. Budget rules, human resource practices, and operational procedures generate powerful incentives to maintain the status quo, and therefore limit organizational learning.

Organizational change is thus a challenging process. Obstacles to change become embedded in the choices organizations make to resolve three dilemmas of collective action: the tensions between policy adherence and continuity and policy adaptation to changed circumstances; between centralized goal setting and decentralized decision making; and between specialization and openness to diverse disciplines.

While powerful corporate oversight mechanisms are needed in well-run organizations they often inhibit nimble adaptation to change. Command and control systems can undercut creativity, innovation, or responsiveness to stakeholders' highly differentiated and evolving needs. Mandated business processes that leave little scope for individual initiative discourage corporate learning and innovation. To be sure control systems are needed because without them, policies and standards may be ignored. In organizations, especially large ones, free-riding behavior is a rational choice for individual group members (Olson, 1971), and can only be contained by countervailing incentives that encourage disciplined effort toward corporate goals and penalize opportunistic behavior. Hence corporate control systems are essential to help secure alignment of staff actions with corporate goals. But especially in large organizations such systems can contribute to rigid decision making. In particular, tough quality-control systems exercised from the top of the hierarchy may have the unintended effect of generating fear of retribution and its unpleasant corollary: the hiding of errors.

Another dilemma of collective action has to do with the obstacles inherent in asymmetrical access to information. The sheer volume of information that must be processed for effective and timely decision making requires delegation of authority. In turn, this raises principal-agent and coordination problems that cannot be resolved without behavioral protocols designed to minimize internal transaction costs.

A third dilemma is associated with the transformation of information management into knowledge. It requires the exercise of professional judgment so that in an organization entrusted with a complex mandate, staff specialization is imperative. But while specialization favors efficiency, it can lead to silo thinking and inhibit the lateral thinking that may hold the key to success when flexibility and creativity are at a premium. Generality of roles promotes flexibility and interdependence. On the other hand tolerance of ambiguity may lead to sloppy decision making.

Sound governance structures, skilled board oversight, and civil society scrutiny can help in ensuring organizational learning and to counter managerial temptations, whether conscious or unconscious, to cover up mistakes, protect the status quo and avoid corrective action. It follows that without independent scrutiny it is hard for the supreme governance authority of a large and complex public organization, let alone the citizenry, to find out what is really going on. Once again, independent evaluation emerges as a critical instrument of managerial accountability.

These then are the organizational tensions that underlie the need for mechanisms designed explicitly to nurture corporate innovation and learning. Evaluation embedded in business processes and focused on results can help to resolve the dilemmas of collective action that plague large, hierarchical organizations.

Promoting Organizational Learning Through Focusing on Results

To help an organization adapt to an evolving environment, its controls must be judiciously designed and its organizational rules periodically adjusted. Quality assurance is needed for organizational learning, but it should be delivered in "real time" and be designed to overcome information asymmetries at affordable cost. It should also be sensitive to changes in the authorizing and operating environment, because fixed standards and rigid bureaucratic norms can undermine organizational responsiveness. Especially in volatile operating environments, spirited debate within the organization, sensitivity to stakeholders' needs and constraints, openness to new ideas, readiness to drop outdated strategies, nimble execution, innovative solutions, and prudent risk management (rather than risk avoidance) are more important than compliance with established policies and rules.

These needs imply that operational protocols and quality assurance standards should be kept under continuous review through evaluative processes. While corporate oversight mechanisms tend to inhibit timely adaptation to change, organizational learning facilitates change by influencing how the organization sorts, processes, stores, and uses the information it needs for decision making.

Evaluation, especially in large organizations, helps shift incentives for management to a much needed counterweight within the corporate incentive structure by shifting the focus of corporate management from inputs to results.

Utilization of Evaluation Lessons: Organizational and Cultural Context Is Key

How does evaluation induce positive organizational change and learning? High-quality and independent reports do not guarantee the use of lessons drawn. The influence of an individual evaluation report is affected by many other factors. Relevance, timing, and dissemination methods all play an important role. But the absorption of evaluation lessons is typically subject to complex political and administrative dynamics (Weiss, 1998). What matters most in evaluation use is the organizational and cultural context within which evaluation is conducted and whether this context favors organizational learning.

Therefore, independent evaluation should not be judged solely by its results. Independent evaluators should plan, design, and disseminate their products so as to facilitate evaluation use. But to ensure objectivity, they must operate at arm's length from decision making. Hence they cannot be held accountable for the utilization of their recommendations: this responsibility lies squarely with policymakers and program managers.

Furthermore, independent evaluation may pay its way even where the internal organizational culture is defensive and averse to change, that is, where the *instrumental use* of evaluation—the straightforward application of valid lessons learned from past programs in the design and implementation of new programs—is limited. This is so, because the evaluation process is apt to generate complex, subtle, and frequently delayed reactions.

Where independent evaluation is resisted, countervailing control mechanisms that seek to elicit management responses to specific evaluation recommendations in a linear fashion tend to be ineffective, at least in the short run. Nonetheless, findings that are brushed aside in the first instance may trigger internal debate, generate stakeholder pressures, and induce public scrutiny that in time may evoke positive change. Gradual, begrudging, and tacit acceptance of evaluation prescriptions is not uncommon. Partial and sometimes hidden reforms may result, for example when change agents within the organization decide to take action within their own sphere of influence. In particular ideas generated by an evaluation may fall on fertile ground following a crisis or when a new management has taken the reins of the organization. As a result, higher-order policy changes involving diverse and powerful interests may be induced. A few such cases suffice to justify corporate investments in evaluation.

Furthermore, instrumental use in the short or medium term is only one of the potential benefits of evaluation. Instrumental use has obvious advantages, but these benefits cannot be fully tapped unless evaluation becomes embedded in the organizational culture. In an organization resistant to learning, *conceptual use* of evaluation findings may be more significant than instrumental use. A sound evaluation process may in time influence program staff to sharpen policy and program design and it may promote ideas that are likely to improve implementation. Evaluation may also empower internal change agents, by confirming their insights or by bringing to light inconvenient realities that had previously been swept under the rug. Finally, evaluation has public good characteristics in that it may have positive effects outside the organization through knowledge creation and contributions to public understanding.

A valid rationale for investing in independent evaluation is that achieving timely organizational adjustments in a turbulent and demanding environment should not be left to chance. Past organizational achievements do not necessarily presage success. In fact organizational success sometimes leads to complacency and resistance to change.

The Independent/Self-Evaluation Nexus

From an organizational learning perspective, the core challenge of independent evaluation is to influence corporate and operational functions while operating at arm's length from them. Precisely because independent evaluation is selective and intermittent it needs to leverage self-evaluation in order to be most effective. The right balance between self- and independent evaluation should be guided by the principle of "subsidiarity," stepping aside from domains that are best handled by self-evaluation and focusing on higher-level questions that are not adequately assessed by self-evaluation. Independent evaluation units that are fully shielded from vested interests are better placed to produce reliable and uncompromising reports. Findings of self-evaluation are more likely to be owned and implemented by decision makers since they are self-generated. Self-evaluation has lower information-generation costs than independent evaluation, and it can help improve program and policy design in a timelier manner.

Independent Evaluation Needs Self-Evaluation

To reach its full potential, independent evaluation needs to leverage its impact through self-evaluation. Self-evaluation can and should serve as one of the transmission belts that connect independent evaluation findings to management decisions and the operational and policy cycle. Often, self-evaluation can address issues that are still subjected to "cooling-off periods" for independent evaluation, identifying corrective actions and providing assessments that can eventually serve as the basis for an independent review.

Self-Evaluation Needs Independent Evaluation

Independent oversight makes self-evaluation more effective. Independent evaluators prod self-evaluators to be more skeptical and reflective about their assumptions, preconceptions, and interests. The mindset of independent evaluators induces self-evaluators to think harder about what the organization is trying to accomplish, to consult more systematically with stakeholders, and to achieve a more resilient consensus about program goals. Independent evaluation also safeguards accountability if self-evaluation is weak.

In the words of Michael Scriven, independent evaluation "can decrease certain types of bias (including) . . . extreme conflicts of interest where the evaluator is 'in bed with' the program being evaluated . . . typical of much program monitoring by agencies and foundations where the monitor is usually the godfather of the program, sometimes its inventor, and nearly always its advocate at the agency" (Scriven, 1991).

Self-Evaluation and Independent Evaluation Should Be Closely Connected

A combination of independent and self-evaluation encourages managers to design evaluable programs—that is, programs with clear goals, verifiable objectives, and adequately funded monitoring and evaluation arrangements that are built upfront into the program design. Within such a framework, self-evaluation (and auditing) focus on compliance and "doing things right," while independent evaluation mostly focuses on "doing the right things" as well as on quality assurance of self-evaluation and validation of its results.[5]

Right-Sizing the Evaluation Function

Sufficient budgets and skills should be allocated to both independent evaluation and self-evaluation in order to allow a fair, accurate, and well-documented assessment of overall organizational effectiveness. A critical mass of resources allocated to both functions is needed for authority to be held responsible. Without adequate budgets, evaluation risks being relegated to a symbolic role.[6]

Self-evaluation should have sufficient resources to ensure that real-time monitoring and lessons learned contribute to program design and corrective actions. Independent evaluation should receive sufficient resources to address the high-level questions that self-evaluation is unlikely to ask, as well as to spot-check the validity of self-evaluation claims.

The Critical Role of Independent Evaluation in the IFIs

Why are the above considerations especially relevant to international financial institutions? The main reason is that rightly or wrongly IFIs are widely perceived to be insufficiently effective and accountable (Woods, 2001). The chronic instability of the international economy and the stubborn persistence of global poverty underlie public dissatisfaction with their performance, and thus they are and are likely to remain under intense public scrutiny. As intergovernmental bodies they are not directly accountable to citizens through the ballot box. Instead, they are governed by executive boards on which their member countries are represented—but with sharply different voting powers based on formulas that ignore population sizes and have yet to be adjusted to reflect adequately the current relative weights of individual countries in the global economy.

The IFIs' legitimacy rests on the degree to which their executive boards make decisions that comply with the agreed mandates of the institutions, promote the general welfare, and are free from political interference. In turn, the influence actually exercised by the executive boards is circumscribed by the considerable power delegated to their chief executives, who control day-to-day operations and whose selection has traditionally been controlled by some of the largest shareholders.

This governance structure implies a democratic deficit that can only be filled, albeit imperfectly, through the pressure of a global public opinion shaped by civil

---

[5]The contributions of self- and independent evaluation to organizational learning can also be expressed in terms of *single* and *double-loop learning*, concepts from the organizational science literature (Argyris, 1977).

[6]Equally, the evaluation function should not become too big. Beyond a certain point, diminishing returns set in and the evaluation function may add to transaction costs without inducing much additional accountability or learning. In Kenneth Arrow's terms, "To serve its functions, responsibility must be capable of correcting errors but should not be such as to destroy the genuine values of authority. Clearly, a sufficiently strict and continuous organ of responsibility can easily amount to a denial of authority. . . . To maintain the value of authority, it would appear that responsibility must be intermittent." See Arrow, 1974, pp. 77–78.

society groups, academic institutions, private organizations, the mass media, social networks, etc.

In this contested public space, independent evaluation reports evince far more interest than self-evaluation reports produced by management. Internally, too, independent evaluation is influential since it reports to the owners of the institution through the board. This reporting relationship matters a great deal in the highly hierarchical, technocratic, elitist, and inward-oriented cultures that frequently characterize the IFIs.

Finally, it is noteworthy that the combination of a strong executive and a highly skeptical public opinion has led to the design of complex and cumbersome internal management processes heavy on controls and characterized by limited tolerance for deviations from norms—that is, an organizational culture that prizes compliance. Here again, the role of independent evaluation comes into sharp focus, given its critical contribution to rethinking of goals and established policies, processes, and practices.

Conclusions

Organizational learning differs from individual learning. Beyond knowledge acquisition, the organizational learning process requires attitudinal shifts and behavioral changes. To achieve organizational effectiveness, staff need to be induced to "do things right"—that is, to observe currently agreed procedures and focus on corporate goals. But for an organization to continue "doing the right things" in a changing environment also requires periodic changes in the strategies, policies, and processes that shape organizational behavior.

These changes are made easier with the help of high-quality evaluation that assesses the validity of established strategic objectives and reconsiders the rationale of ingrained business processes. Ensuring compliance and that organizations "do things right" is mostly the province of monitoring and auditing. Identifying the "right things to do" and inducing organizations and their staff to move in those new directions is mainly the province of independent evaluation.

Independence is an essential ingredient of evaluation excellence. Independent evaluation helps to enhance the quality and credibility of evaluation products and contributes to organizational transparency and accountability. Along with quality assurance and ethical guidelines, evaluation independence is a widely recognized feature of organizational credibility and reliability. While self-evaluation is closer to the action and thus can more rapidly impact organizational behavior, it often lacks the distance needed to achieve objective assessment. Therefore, independent evaluation should assess the validity of self-evaluation findings just as external auditors assess the validity of internal accounting.

The benefits generated by independent evaluation are a function of the organizational context. They can be very large in open and accountable working environments where innovation and creativity are rewarded. In such environments, evaluation independence induces public confidence, protects the learning process, and induces program managers and stakeholders to focus on results.

To be most effective, independent evaluation needs to be appropriately connected to the rest of the organization. Independence should not be confused with isolation. Indeed, evaluation does not facilitate organizational learning if it fails to feed into strategy formulation, to amplify the voice of legitimate stakeholders, and/or to provide credible and reliable performance information to management and higher governance authorities. In particular, independent evaluation needs to leverage its impact through appropriate links to self-evaluation processes, including judicious protocols of professional interaction. To achieve and maintain such relationships calls for careful institutional design of organizational structures and business processes.

Finally, independent evaluators need to be adequately protected to deliver high-quality, uncompromising reports and to shield the function from capture and intimidation. For evaluation to be genuinely independent, all the essential enabling conditions—structural, behavioral, protection from external influences, and avoidance of conflict of interest—need to be in place. Both independent and self-evaluation should be adequately resourced to have a material impact on the organization and to avoid the perception that they are only there for window dressing.

These conclusions are especially relevant to the IFIs in their unfinished journey towards greater accountability, transparency, and legitimacy. In these organizations, independent evaluation makes a unique and essential contribution to organizational accountability and learning. It does so not only by assessing the validity of self-evaluation findings but also by looking deeper and more objectively at sensitive corporate management issues that self-evaluation is unable to tackle. The comparative advantage of independent evaluation lies in the rigorous and fair assessment of corporate strategies, structures, and processes. In this way independent evaluation helps the organization navigate the turbulent waters of the global economic environment.

Annex. Template for Assessing the Independence of Evaluation Organizations

Criterion	Aspects	Indicators
I. Organizational independence	The structure and role of evaluation unit.	Whether the evaluation unit has a mandate statement that makes clear its scope of responsibility extends to all operations of the organization, and that its reporting line, staff, budget, and functions are organizationally independent from the organization's operational, policy, and strategy departments and related decision making.
	The unit is accountable to, and reports evaluation results to, the head or deputy head of the organization or its governing board.	Whether there is a direct reporting relationship between the unit, and (a) the management and/or (b) board or (c) relevant board committee, of the institution.
	The unit is located organizationally outside the staff or line management function of the program, activity, or entity being evaluated.	The unit's position in the organization relative to the program, activity, or entity being evaluated.

Annex (*continued*)

Criterion	Aspects	Indicators
	The unit reports regularly to the larger organization's audit committee or other oversight body.	Reporting relationship and frequency of reporting to the oversight body.
	The unit is sufficiently removed from political pressures to be able to report findings without fear of repercussions.	Extent to which the evaluation unit and its staff are not accountable to political authorities, and are insulated from participation in political activities.
	Unit staffers are protected by a personnel system in which compensation, training, tenure, and advancement are based on merit.	Extent to which a merit system covering compensation, training, tenure, and advancement is in place and enforced.
	Unit has access to all needed information and information sources.	Extent to which the evaluation unit has access to the organization's (a) staff, records, and project sites; (b) co-financiers and other partners, clients; and (c) programs, activities, or entities it funds or sponsors.
II. Behavioral independence	Ability and willingness to issue strong, high-quality, and uncompromising reports.	Extent to which the evaluation unit: (a) has issued high-quality reports that invite public scrutiny (within appropriate safeguards to protect confidential or proprietary information and to mitigate institutional risk) of the lessons from the organization's programs and activities; (b) proposes standards for performance that are in advance of those in current use by the organization; and (c) critiques the outcomes of the organization's programs, activities, and entities.
	Ability to report candidly.	Extent to which the organization's mandate provides that the evaluation unit transmits its reports to the management/board after review and comment by relevant corporate units but without management-imposed restrictions on their scope and comments.
	Transparency in the reporting of evaluation findings.	Extent to which the organization's disclosure rules permit the evaluation unit to report significant findings to concerned stakeholders, both internal and external (within appropriate safeguards to protect confidential or proprietary information and to mitigate institutional risk). Who determines evaluation unit's disclosure policy and procedures: board, relevant committee, or management?

(continued)

Annex (*continued*)

Criterion	Aspects	Indicators
	Self-selection of items for work program.	Procedures for selection of work program items are chosen, through systematic or purposive means, by the evaluation organization; consultation on work program with management and board.
	Protection of administrative budget, and other budget sources, for evaluation function.	Line item of administrative budget for evaluation determined in accordance with a clear policy parameter, and preserved at an indicated level or proportion; access to additional sources of funding with only formal review of content of submissions.
III. Protection from outside interference	Proper design and execution of an evaluation.	Extent to which the evaluation unit is able to determine the design, scope, timing, and conduct of evaluations without management interference.
	Evaluation study funding.	Extent to which the evaluation unit is unimpeded by restrictions on funds or other resources that would adversely affect its ability to carry out its responsibilities.
	Judgments made by the evaluators.	Extent to which the evaluator's judgment as to the appropriate content of a report is not subject to overruling or influence by an external authority.
	Evaluation unit head hiring/firing, term of office, performance review and compensation.	Mandate or equivalent document specifies procedures for the (a) hiring, firing, (b) term of office, (c) performance review, and (d) compensation of the evaluation unit head that ensure independence from operational management.
	Staff hiring, promotion, or firing.	Extent to which the evaluation unit has control over: (a) staff hiring, (b) promotion, pay increases, and (c) firing, within a merit system.
	Continued staff employment.	Extent to which the evaluator's continued employment is based only on reasons related to job performance, competency, or the need for evaluator services.
IV. Avoidance of conflicts of interest	Official, professional, personal, or financial relationships that might cause an evaluator to limit the extent of an inquiry, limit disclosure, or weaken or slant findings.	Extent to which there are policies and procedures in place to identify evaluator relationships that might interfere with the independence of the evaluation; these policies and procedures are communicated to staff through training and other means; and they are enforced.

Annex (*continued*)

Criterion	Aspects	Indicators
	Preconceived ideas, prejudices, or social/political biases that could affect evaluation findings.	(a) Extent to which policies and procedures are in place and enforced that require evaluators to assess and report personal prejudices or biases that could imperil their ability to bring objectivity to the evaluation; and (b) Extent to which stakeholders are consulted as part of the evaluation process to ensure against evaluator bias.
	Current or previous involvement with a program, activity, or entity being evaluated at a decision-making level, or in a financial management or accounting role; or seeking employment with such a program, activity, or entity while conducting the evaluation.	Extent to which rules or staffing procedures that prevent staff from evaluating programs, activities, or entities for which they have or had decision-making or financial management roles, or with which they are seeking employment, are present and enforced.
	Financial interest in the program, activity, or entity being evaluated.	Extent to which rules or staffing procedures are in place and enforced to prevent staff from evaluating programs, activities, or entities in which they have a financial interest.
	Immediate or close family member is involved in or is in a position to exert direct and significant influence over the program, activity, or entity being evaluated.	Extent to which rules or staffing procedures are in place and enforced to prevent staff from evaluating programs, activities, or entities in which family members have influence.

Sources: U.S. General Accounting Office, *Government Auditing Standards, Amendment 3* (2002); OECD/DAC Working Party on Aid Evaluation, *Glossary of Key Terms in Evaluation and Results Based Management* (2002); OECD/DAC, *Principles for Evaluation of Development Assistance* (1991); INTOSAI, *Code of Ethics and Auditing Standards* (2001); Institute of Internal Auditors, *Professional Practices Framework* (2000); European Federation of Accountants, *The Conceptual Approach to Protecting Auditor Independence* (2001); Danish Ministry of Foreign Affairs, *Evaluation Guidelines* (1999); and Canadian International Development Agency, *CIDA Evaluation Guide* (2000). See https://wpqr1.adb.org/LotusQuickr/ecg/Main.nsf/h_9BD8546FB7A652C948257731002A062B/AA95B62CF943F82E4825774B003A82E0/?OpenDocument&Form=h_PageUI.

References

Argyris, Chris, 1977, "Double-loop learning in organizations," *Harvard Business Review.* Available at www.westernsnowandice.com/09-Presos/DoubleLoop.pdf.

Arrow, Kenneth J., 1974, *The Limits of Organization* (New York: W.W. Norton).

General Accountability Office of the United States (GAO), *Government Auditing Standards.* See www.gao.gov/yellowbook.

International Federation of Accountants, 2010, *Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Service Pronouncements* (New York).

Mayne, John, 2008, *Building an Evaluative Culture for Effective Evaluation and Results Management*, ILAC Brief 20 (November). Available at www.cgiar-ilac.org/files/ILAC_Brief20_Evaluative_Culture.pdf.

Olson, Mancur, Jr., 1971, *The Logic of Collective Action: Public Goods and the Theory of Groups* (Cambridge, Massachusetts: Harvard University Press).

Scriven, Michael, 1991, *Evaluation Thesaurus* (Newbury Park, London, and New Delhi: Sage Publications, 4th ed.).

Weiss, Carol H., 1998, "Have We learned anything new about the use of evaluation?" *American Journal of Evaluation*, Vol. 19, No. 1, pp. 21–33.

Woods, Ngaire, 2001, *The International Monetary Fund and the World Bank*. Available at www.globaleconomicgovernance.org/wp-content/uploads/Woods%20for%20Routledge.pdf.