Journal of MultiDisciplinary Evaluation
Volume 9, Issue 20, 2013

ISSN 1556-8180
http://www.jmde.com

Evaluation Independence In Organizations

Robert Picciotto
Kings College, London

Background: Evaluation independence is a neglected issue in the literature. Yet, it is a critical aspect of evaluation governance in organizations.

Purpose: The article draws on organizational theory, institutional economics and international development evaluation practice to define evaluation independence in organizations, outline principles geared to the design of evaluation processes within organizations and trigger a debate on evaluation independence in the evaluation community.

Intervention: Not applicable.

Research Design: Literature review informed by personal exposure to management of the World Bank Independent Evaluation Group.

Data Collection and Analysis: Not applicable.

Findings: Independent and self-evaluation are complementary and synergistic. Organizational theory demonstrates that a judicious combination of independent and self-evaluation contributes to managerial accountability and double loop learning.

Keywords: *independence; accountability; internal evaluation; external evaluation; double loop learning*

Introduction

"There is much to be done in the design of institutions to reconcile the values of authority and responsibility" (Arrow, 1974, p. 77). Yet, evaluation independence is a neglected topic in the literature. To help fill the gap this article puts forward a normative framework about the role of independent and self-evaluation in organizations. It draws on organizational theory as well as on the author's own experience at the helm of the World Bank's Independent Evaluation Group (1992-2002). What is independent evaluation? What is its rationale? 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?

Background

Business school gurus and institutional economists have yet to catch up with the growing role of evaluation in organizations. Within the evaluation community, the governance dimension of evaluation is rarely addressed and evaluation independence is frequently viewed as an ethical standard that evaluators should comply with in order to generate unbiased analyses and credible findings.

This narrow and individualistic conception of evaluation independence has consequences. It precludes scrutiny of functional and structural independence and it ignores the overt and covert pressures that evaluators are routinely subjected to in the real world. With rare exceptions, the role of commissioners is not recognized as a critical determinant of evaluation quality in evaluation guidelines (Picciotto, 2005). Instead, the integrity of evaluations is perceived to rest primarily on the competencies of evaluators, the quality assurance arrangements in place, the appropriateness of evaluation methods and the transparency of evaluation processes.

The neglect of evaluation independence as a public and corporate management concern is consistent with the widespread view that external evaluations are independent. Yet more often than not fee dependence and contractual constraints severely restrict the freedom of external evaluators to exercise unbiased judgment. Equally, the lack of scholarly attention to evaluation independence may be due to the mistaken notion that evaluation, just as accounting and auditing, has achieved the status of fully fledged profession which enjoys autonomous control over its occupational practices. Unfortunately, this is not the case (Picciotto, 2011).

Finally, the evidence base about the relative effectiveness of independent evaluation vs. self-evaluation is thin except in the international development domain. Elsewhere, few public, private and voluntary sector organizations have equipped themselves with internal and independent evaluation functions. Yet there is palpable public disenchantment with external evaluations that are not truly independent (Love, 1991), while internal evaluation arrangements are patently inappropriate for accountability purposes given the pervasiveness of self-regarding behaviour in organizations and its corollary: a reluctance to accept failure (Vedung, 1997).

Looking ahead, it is likely that independent evaluation in organizations will come into the spotlight given the lasting preoccupation with public sector reform. The lofty rhetoric of results that induced the *new public management* movement to promise utopian outcomes through devolution, decentralization and deregulation has given way to a realistic stance that recognizes the limits of market based solutions, the diversity of organizational contexts and the value of incremental improvements through internal adjustment mechanisms (Pollitt & Boukaert, 2000).

What is Evaluation Independence?

Ambivalence about evaluation independence is widespread among practitioners. Only evaluators who subscribe to democratic evaluation principles are strong proponents of evaluation independence. They are a minority. Most evaluators operate as consultants. They give primacy to the utilization of evaluation results and limit their ambition to the timely provision of evaluative knowledge to decision-makers. They are also concerned about the chilling effect and the isolation that independence may induce.

As a result these evaluators favor a strictly advisory supporting role that does not challenge the implicit values of their employers or the processes that underlie organizational performance[1]. To be sure, evaluation independence is not an end in itself. But, as the rest of this article seeks to demonstrate, the logic of evaluation independence is rooted in organizational theory (Arrow, 1974). When all is said and done, organizations exist to resolve the inherent tensions that exist between individual and collective goals; to manage information flows; and to coordinate actions through a nexus of contracts that keeps transaction costs in check. Independent evaluation has a key role to play in all of these functions.

Enforcement of codes of conduct within organizations implies authority and, except in dictatorial regimes which brook no opposition, authority is circumscribed and managerial freedom of action is not absolute. This is because authority is poorly tolerated and cannot be sustained over time unless it is legitimate. In turn, legitimacy implies that the persons entrusted with running the organization are both responsible and accountable to a higher authority and the wider public. Hence, holding 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 the risks could have been managed differently so that the errors were (with the benefit of hindsight) unnecessary.

Similar mechanisms help to inform the adaptation of internal processes and protocols through organizational learning. Of course, errors must be avoided in the evaluative process itself and this means that evaluation processes should be designed and managed so as to assess the merit and worth of organizational activities in a fair, valid and accurate fashion. These then are the imperatives of sound governance that evaluation independence helps to sustain.

Independent evaluation helps to protect the integrity of the management process. It enhances its credibility and provides fresh perspectives on the policies and programs being evaluated (Mayne, 2008). These goals cannot be met if evaluations are contracted in ways that:

- constrain information so that evaluation products cannot have any critical content;

- control the content of the evaluation program so that it does not contribute new knowledge;
- delay the evaluation process (or the disclosure of evaluation results) until after the decisions that might have been informed by the evaluation are taken; and
- induce evaluators to focus on irrelevant or marginal aspects of the program or policy being evaluated.

Independence and Quality are Synergistic

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, pp. 192-193).

This said, independence on its own does not guarantee evaluation quality: relevant skills, sound methods, adequate resources and transparency are also required. Independent but incompetent evaluations can be misleading, disruptive and costly. On the other hand, evaluation quality without independence lacks credibility. This is why persons and entities that have reason to fear the outcome of an evaluation will frequently throw doubt on its independence.

In open and accountable working environments, evaluation independence induces confidence, protects the learning process and induces program managers and stakeholders to focus on results. Along with quality assurance and ethical guidelines, independence properly defined should be a recognized feature of credibility and reliability for internal evaluation systems and this should be reflected in the evaluation guidelines issued by evaluation associations.

Independence is a Competency

Evaluation independence is primarily a characteristic of corporate governance. But it does have implications for evaluators' competencies. In evaluation as in science, knowledge and practice are not enough to achieve excellence. Dispositions and attitudes matter as well. They are characterized by curiosity, skepticism and a hunger for evidence. In addition evaluation calls for courage and fairness reflected in a balanced and self-confident approach that does not shirk from pointing out problems and performance shortfalls but also recognizes success and achievement. Thus the Yellow Book of the General Accountability Office of the United States that sets out criteria for auditing and evaluation identifies "an independent attitude and appearance" as desirable characteristics.

Similarly, according to the Handbook of the International Federation of Accountants (2003) independence of mind "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" while independence in appearance is "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 scepticism had been compromised"). Both meanings are duly recognized as evaluators' capabilities by the European Evaluation Society (2011).

Independent Evaluators Need Protection

Independent dispositions matter but one cannot reasonably expect evaluators to act independently in organizational contexts that put their careers, their reputations and their livelihoods at risk. They need adequate protection to carry out their demanding and stressful mandate. Special institutional safeguards are needed. Specifically, evaluators need to be shielded from external threats to their impartiality. They should be given full access to the information they need to carry out their work. Immunity from capture by any of the parties that share program management responsibility is fundamental.

How is Evaluation Independence Granted and Protected?

If external evaluators are commissioned to carry out an evaluation the process is governed by a contract. The work is also guided by voluntary guidelines issued by evaluation associations. But the actual 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 whether he/she is genuinely trying to find out whether a policy or a program works and if not why and how it might be improved.

Thus evaluation independence is a "hit and miss" affair in circumstances where the organization that contracts an evaluation has not specified precise safeguards that guarantee independence and transparency for the evaluation process. What are these safeguards? They have been codified for cases where the evaluation is carried out or contracted by a unit embedded within an organization. Specifically, four major dimensions of evaluation independence have been identified by the Evaluation Cooperation Group (2010) based on a careful literature review and decades of independent evaluation experience in the international development domain: (a) organizational independence; (b) behavioral independence; (c) avoidance of conflicts of interest; and (d) protection from external influence. Additional information is provided in table 1.

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 fulfil 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 to 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. All of them matter. Protection from outside interference is the ultimate aim of organizational independence. Conflicts of interest are frequent absent organizational independence. Behavioral independence is a function of organizational independence as well as avoidance of conflicts of interest and protection from external interference[2].

Externality does not Guarantee Independence

It should be clear by now that external evaluation cannot be equated with evaluation independence. Fee dependence is a major threat to the integrity of the evaluation process. The judgment of external evaluators is impaired or threatened if their services are retained by the managers in charge of the activities being evaluated. By contrast, internal evaluation units funded and controlled by the supreme governance authority are better protected from management interference. This is because of the checks and balances associated with the statutory oversight role that boards of directors are tasked to play.

At the same time internal independent evaluation units backed up by an adequate mandate and staffed with evaluators familiar with the organizational terrain enjoy relative proximity with the programs being evaluated. Thus evaluations carried out by internal independent units reporting to boards of directors are far more likely to overcome "information asymmetries" while protecting the objectivity of the evaluative process.

Independence is not Isolation

The acid tests of independence and quality have to do with the extent to which evaluations are endowed with:

- *criticality:* the ability and willingness to judge performance in an objective and transparent fashion;
- *additionality*: a distinctive contribution to operational knowledge creation or dissemination;
- *timeliness:* the delivery of operations evaluation findings and lessons early enough to inform decision making; and
- *materiality*: a deliberate focus on topics and issues that have substantial relevance to development effectiveness.

Some evaluators argue cogently that having no connection or shared experience with intended users of evaluation findings constrains access to information, evinces resistance and inhibits learning. There is little doubt that accurate and fair evaluations combine intellectual detachment with empathy and understanding. The ability to engage with diverse stakeholders and secure their trust while maintaining the integrity of the evaluation process is a mark of evaluation professionalism. Contestability of independent evaluation findings and principled, respectful disagreements about recommendations make for a healthy organizational learning 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 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.

Independent Evaluation can Help Connect Program Knowledge with Strategy

Organizational theory holds that the creation of public value by an organization requires operational capabilities that are aligned with the demands of its *authorizing environment (Moore, 1995)*. Hence, meeting the accountability and knowledge needs of stakeholders is central to the design of a responsive evaluation function. It is especially important in the public and voluntary sectors since the main restoration mechanism to poor performance in the public sphere is the "voice" option whereas in the private sector the "exit" option exercised through competitive market mechanisms is how stakeholders express their dissatisfaction (Hirschman, 1970).

Thus disgruntled shareholders can sell their shares and unhappy consumers can shift from one brand or supplier to another. Another difference between the economic and political market place is that private suppliers serve fairly homogeneous needs that are readily translated into measures of merit and worth. Specifically private corporations have ready access to information about their performance (i.e. market tests and financial ratios) whereas, in the public sphere, social services meet varied interests that must be satisfied from one electoral cycle to the next.

In the interval between electoral cycles, feedback mechanisms must take account of the views of diverse stakeholders and strike appropriate tradeoffs among them. Public action impacts and works through the private sector and the civil society. This means that policy makers and managers of social programs cannot create value through exclusive reliance on factors which they can control fully or directly. In particular, public opinion is the life blood of government. For public sector organizations this puts special emphasis on community relations and information campaigns that can trigger voluntary action and public support. For evaluation, therefore, the *external operating environment* should help generate the indicators that inform strategic decisions. Accordingly, an independent evaluation function is well placed to validate indicators and processes that connect public and voluntary sector organizations with the citizens they serve. Feedback from diverse constituencies is secured through a variety of time honored participatory evaluation instruments (focus groups, surveys, etc.) and increasingly through social media technologies.

In sum evaluation cannot fulfill its potential unless it connects effectively to its management, the supreme authorities that govern the organization and the broader society. But in doing so it should maintain its objectivity, exercise full freedom of inquiry and resist capture. This is imperative since without independence evaluation is perceived by the public to be subservient to vested interests: for evaluation knowledge to be credible, legitimate and valuable evaluation must be functionally and structurally independent.

Independent Evaluation Enhances Organizational Accountability

Managers whether in the private or public sector are mandated to secure high value for the bundle of assets assigned to their care (Moore, 1995). They need to demonstrate to those in control of their funding that they are managing the resources entrusted to them responsibly and effectively. They are accountable to deliver results. But public and private organizations secure the funds needed to carry out their activities in different ways. In the private sector, revenues derive from the sale of goods and services to individual consumers – or from the private capital markets that judge the prospective value of such sales.

In the public sector the ultimate 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 judiciously targeted to persons in authority. The difficulty of measuring social value explains why such simple and inexpensive output measures and budget coefficients rather than outcomes and impacts have traditionally dominated public sector management.

Such indicators 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 their work needs independent validation: independent evaluation in the public sector is what auditing of accounts is in the private sector. This is why the need for a close linkage between independent evaluation and organizational accountability is self-evident. 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.

What is Organizational Learning?

In a learning organization, employees are engaged in lifelong learning. Learning can only be ascertained with confidence through changes in behavior. But not all behavioral changes result from knowledge acquisition or experience. Conditioning through fear and reward, respect for authority, desire to conform, etc. also affect individual actions. Learning only takes place if one's thoughts and actions have been shaped by a fresh understanding of the environment in which one lives.

Learning is not automatic. External intervention is critical for self improvement and deep understanding. Learning challenges current ideas and preconceptions. It may be triggered by unexpected events or external threats. It may also be nurtured through coaching, training or formal education. It leads to new insights and encourages acquisition of new concepts and mental models that serve as useful guides to action. This has important implications for organizational effectiveness.

Without appropriate knowledge, experience and skills individuals left to their own devices exhibit persistent shortfalls in understanding grounded in misconceptions, overgeneralization, mistaken interpretation of evidence, inability to formulate cogent viewpoints, etc. (Perkins, 1993). Cognitive science, educational psychology, and practical experience demonstrate that learning is not automatic and requires interaction with experienced, knowledgeable, external agents. This is why organizational effectiveness depends in significant part on sound human resource policies, a critical dimension of organizational effectiveness…but this is not the end of the story.

Organizational Learning has Its Own Dynamics

Public policies and programs are designed and implemented by and within organizations. Decision makers within them are not free of constraints. They are conditioned to comply with "rules of the game". These are shaped by legal, procedural and traditional constraints and by time honored customs that favor habitual ways of doing things. Budget rules, human resource practices, operational procedures generate powerful incentives.

Hence, organizational learning is not the same as individual learning. To be sure, it cannot be divorced from individual learning since individuals are what make an organization tick. But most important organizational and strategic decisions are collective decisions shaped by hierarchy, protocols and precedent. Through independent evaluation organizational learning and individual learning are bridged. This is why the widely held proposition that independence in evaluation is not conducive to learning is wrong. Well governed organizations treat organizational learning as a managerial accountability.

Dilemmas of Collective Action

In organizations, especially large ones, free riding behavior is a rational choice for individual group members (Olson, 1971). This problem can only alleviated by countervailing incentives that encourage disciplined effort towards corporate goals and penalize opportunistic behavior. This explains why control systems are needed in well run organizations. Without them, policies and standards may be ignored.

Equally, organizations must overcome 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 'rules of the games' designed to minimize internal transaction costs. Command and control systems are needed but when they are tight and rigid they undercut creativity, innovation or responsiveness to highly differentiated stakeholders' needs. Accordingly periodic changes in the control environment are needed to help the organization adapt to an evolving environment. Once again this is where independent evaluation comes in.

Enter Evaluation

The path dependence commonly associated with fixed standards and bureaucratic norms undermines organizational responsiveness. Quality assurance is therefore needed for organizational learning but it should be delivered in "real time"; it should be designed to overcome information asymmetries at affordable cost and it should be sensitive to changes in the authorizing and operating environment. Hence, operational protocols and quality assurance standards should be kept under continuous review through *evaluative* processes.

In sum, corporate control systems are essential to help secure alignment of staff actions with corporate goals, especially in large organizations. But they can also contribute to slow, costly and 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. Thus detailed business protocols may have an unintended consequence: limited corporate learning and innovation. This is fundamentally why independent evaluation is needed to "shake things up" especially in large and successful organizations: an independent evaluation function provides a much needed counterweight within the corporate incentives structure by shifting the focus of corporate management from inputs to results.

The Limits of Charismatic Leadership

To be sure, smart, self-confident, charismatic managers sensitive to the evolving demands of the operating environment can help to minimize the downside risks associated with rigid corporate control systems. But despite their loudly self proclaimed commitment to accountability, transparency and learning such strong leaders tend to be reluctant to acknowledge errors. They frequently equate loyalty to them with loyalty to the organization and they are prone to using the media and modern communications techniques to blunt the bracing effects of unalloyed transparency. Sound governance structures, skilled board oversight and civil society scrutiny can help pierce through the public relations fog that charismatic leaders are apt to create but without independent evaluation 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 and generate a competing narrative.

Evaluation for Organizational Learning

Most organizations still under-invest in evaluation. Their leaders do not recognize the positive role that evaluation can play in improving organizational performance. The reality is that a culture of compliance is not conducive to sound risk management in the real world. Traditional corporate oversight mechanisms have their place but they often inhibit nimble adaptation to change. This is what organizational learning guided by independent evaluation is about. It facilitates change by influencing how the organization sorts, processes, stores and uses the information needed for decision making.

Transforming information into knowledge requires the exercise of professional judgment. If the organizational mandate is complex staff specialization is imperative. But disciplinary biases may distort quality assurance systems and limit innovations. Multiple perspectives and effective internal communications are needed. While specialization favors expertise and efficiency silo thinking can inhibit lateral thinking that may hold the key to success when flexibility and creativity are at a premium. Evaluation is well equipped to address these risks.

Generality of roles promotes flexibility and interdependence but excessive tolerance for ambiguity can lead to sloppy decision making. It follows that learning organizations are those that have mastered the dilemmas of collective action associated with information management and decision making within groups. 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 often more important than compliance with established policies and rules.

Organizational change is therefore a challenging process given the obstacles inherent in the tensions that exist between policy continuity and adaptation to changed circumstances; centralization of goal setting and decentralization of decision making; specialization and openness to diverse disciplines. These then are the paradoxes that underlie the need for explicit mechanisms designed to nurture corporate innovation and learning. Independent evaluation is one of these mechanisms.

Learning from Experience

Embedded in the business processes and focused on results independent evaluation can help to resolve the dilemmas of collective action that plague large, hierarchical organizations. How then does evaluation induce positive organizational change? A vast literature has accumulated on this topic. Unsurprisingly the consensus of opinion is that evaluation quality is an imperative. A misleading and unfair evaluation can cause major organizational damage. The Hippocratic Oath ("First, do no harm") applies to evaluators as it does to doctors. Poor evaluation quality damages the evaluation discipline.

This said, the influence of an individual evaluation report is only loosely and partially correlated with its rigor and quality. Expert and independent distillation of experience does not guarantee utilization of lessons drawn. Other factors (relevance, timing, dissemination methods, etc.) matter more. 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 it is auspicious to organizational learning. Elsewhere the evaluation process is apt to generate complex, subtle and frequently delayed reactions.

This is why independent evaluation should not be judged solely by its instrumental results. To be sure, independent evaluators should design, plan and disseminate their products so as to facilitate evaluation use. But to ensure objectivity, they should operate at arm's length from decision making. Independence means that they cannot be held accountable for the utilization of evaluation findings. This responsibility lies squarely with policy makers and program managers.

Evaluation Enjoys Public Good Characteristics

Where independent evaluation is resisted countervailing control mechanisms that seek to link management actions to specific evaluation recommendations in a linear fashion tend to be ineffective, at least in the short run. But findings brushed aside in the first instance may promote internal debate, intensify stakeholder pressures and induce public scrutiny that in the medium and long term will inevitably induce positive change. Gradual, begrudging and tacit acceptance of evaluation prescriptions is not uncommon.

Partial and sometimes hidden reforms may result, e.g. when change agents within the organization decide to take action within their own sphere of influence. Thus, ideas generated by an evaluation may fall on fertile grounds 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 eventually be induced. Only 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 yielded by evaluation. *Conceptual use* may be more significant than instrumental use in an organization resistant to learning. A sound evaluation process may in time influence program staff to sharpen policy and program design and it may promote ideas 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, evaluations may have positive effects outside the organization through *knowledge creation* and contributions to public understanding.

This said a valid rationale for investing in evaluation lies in the sensible notion that achieving timely organizational adjustments in a turbulent and demanding environment should not be left to chance and serendipity. Instrumental use of evaluation, i.e. the straightforward application of valid lessons learnt from past programs in the design and implementation of new programs, has self-evident advantages. These benefits are tapped in full only if evaluation becomes embedded in the organizational culture.

From Single Loop to Double Loop Learning

The past is not always prologue. Past organizational achievements do not necessarily presage continued success. In fact organizational success can be a curse as it often leads to complacency and resistance to change on the grounds that "if it ain't broke it should not be fixed". In order to master successful adaptation to evolving circumstances learning organizations have: (a) secured the right competencies and nurtured the right skills through relevant individual learning; and (b) resolved constructively the tensions that inevitably arise between organizational continuity and change, flexibility and control, internal cohesion and responsiveness to stakeholders.

For individuals as for organizations, detecting and correcting deviations from norms constitutes *single loop learning*. This dimension of organizational effectiveness focuses on "doing things right" and it is the privileged province of monitoring. On the other hand, conducting assessments of the goals, policies and protocols that underlie current organizational behavior in order to "doing the right things" is *double loop learning* which occurs when error is detected and corrected in ways that involve the modification of an organization's underlying norms, policies or objectives (Argyris, 1977). This vital function is what evaluation is all about in an organization.

On the one hand, single loop learning is of critical importance since no large organization can function efficiently without protocols and rules that minimize internal transaction costs, allow delegation, limit coordination requirements and ensure service quality, timeliness of delivery and responsiveness to clients. Strong morale, cohesive teamwork and institutional loyalty are critical ingredients of organizational excellence.

On the other hand, given the complexity and volatility associated with an increasingly interconnected and turbulent global system corporate strategies and operating principles and standards have short shelf lives. They need to be adapted periodically through double loop learning to reflect changes in the operating and authorizing environments. Thus a judicious balance must be struck between controls (single loop learning) and evaluation (double loop learning).

Independent Evaluation as Double Loop Learning

Single loop learning implemented through auditing and other oversight mechanisms is largely sufficient where organizational objectives are broadly endorsed, well defined, time-bound, measurable and unambiguous and the operating environment is stable. In such circumstances performance models focused on readily verifiable measures provide a clear connection between organizational actions and intended outcomes. In this context traditional management models that emphasize control of the processes that link organizational inputs to outputs, outcomes and impacts are fully justified and monitoring enjoys pride of place in the managerial tool kit.

But more often than not large and complex organizations must satisfy multiple policy objectives and they typically operate in contexts characterized by complexity and instability. Hence double loop learning is also needed and linear models that rely on achieving consistency between pre-determined goals and rigid policy instruments hinder organizational effectiveness. A systems approach geared to nimble and pragmatic management is preferable.

Organizational management approaches (and evaluation models) that focus principally on internal performance and assessment appropriate for situations that require joint actions by diverse partners, principled agreements among them, creative resolutions of tensions among competing policy priorities and appropriate safeguards for the poorest and most vulnerable segments of society.

In such circumstances, evaluation should embrace all partners and periodic goal based evaluations should be complemented by goal free impact evaluations using valuing frameworks that are realistic, ethical and legitimate. In particular, evaluations in order to be ethical should amplify the voices of the weaker party in the relationship. They should rely on verifiable success indicators agreed with all stakeholders.

A global perspective is also needed for an international organization. Interventions principally focused on single country interventions may have unintended consequences if they fail to take account of their potential consequences for the international system.

Striking a Fine Balance

To be sure, exclusive reliance on double loop learning can lead to high internal transaction costs and organizational turmoil. It may induce perpetual ferment within the organization, rapidly changing policy agendas and shifting priorities that eventually undermine organizational effectiveness. On the other hand, single loop learning on its own does not help to resolve the paradoxes inherent in the management of large and complex organizations operating in a rapidly changing context.

This is because compliance with norms is what employees are conditioned to work towards even after the issues these norms were designed to address are no longer fit for purpose. Consequently the relative importance of double loop learning (and therefore evaluation) rises in unstable operating environments subject to rapid change. Giving more emphasis to double loop learning (and evaluation) in organizations traditionally committed to single loop learning is especially hard to induce if they have been successful or if current norms reflect the demands of an authorizing environment that is out of whack with current realities.

More often than not the interests that have empowered corporate units to design organizational norms will do their utmost to defend them and they will equip the organizational hierarchy to enforce them and penalize those that march to a different drummer. In compliance oriented organizations employees who raise fundamental questions about existing policies, protocols and standards do not move up the organization ladder. They are shunned by their colleagues and their behavior may be perceived as disloyal. This explains why whistle blowing is so rare in organizations and also why internal watchdog and independent evaluation units are needed and must protect their members from retribution in order to be effective.

The Independent/Self-Evaluation Nexus

It should be clear by now why single loop learning through monitoring is a continuous process supported by internal control units while double loop learning interventions managed by independent evaluation units need not be as frequent. From an organizational learning perspective the core challenge of independent evaluation is to help influence the organizational adaptation without interfering with management. Independent evaluation must find ways of influencing other corporate and operational functions while operating at arm's length from them. Precisely because independent evaluation is selective and intermittent it needs the leverage of self-evaluation in order to be fully effective. This implies of course that the relationship between the two functions is grounded in mutual respect and professionalism.

Independent Evaluation Needs Self-Evaluation

On the one hand double loop learning should be guaranteed by evaluation units fully shielded from vested interests. They may have a useful impact simply through the generation of valid, reliable and uncompromising reports. On the other hand, such units should not be so far removed from the organization that a climate of distrust and misunderstanding disconnects them fully from the operational and policy cycle. To reach its full potential independent evaluation needs to leverage its impact through self-evaluation. It is the transmission belt that connects evaluation findings to management decisions.

Self-Evaluation Enjoys Unique Advantages

Other things being equal, self-evaluation is more efficient than independent evaluation due to its lower information generation and transfer costs. It is definitely timelier, e.g. especially when cooling periods are imposed on the topics selected for independent review. Self-evaluation generates considerable benefits by improving program and policy design through quality assurance, enhanced monitoring and tracking of results.

Finally, self-evaluation findings are more likely to be owned by decision makers since they are self-generated. Consequently self-evaluation has so many advantages over independent evaluation that it should be adopted and made use of to the maximum extent feasible. But it suffers from its own limitations and it needs independent evaluation in order to work best. The reasons are many.

Self-Evaluation needs independent evaluation

Program managers and policy makers often lack evaluation skills. They may have different interests and concerns than program beneficiaries. They often succumb to leaps of faith that lead to faulty or excessively risky decisions. They may be tempted to select evidence that supports their preconceptions. The mindset of independent evaluators prods self-evaluators to be more skeptical and reflective about the assumptions underlying the design of the policies and programs that they design and implement.

Independent evaluators induce self-evaluators to think harder about what the organization is trying to accomplish, to consult more systematically with stakeholders, to achieve a more resilient consensus about program goals and to mobilize scattered resources and energies towards program goals. Independent evaluation also safeguards accountability when self-evaluation is weak. Professional oversight of self-evaluation by independent evaluation enhances the effectiveness of self-evaluation in ways similar to those that make internal auditing useful in the judicious implementation of administrative policies and procedures.

To be sure, as the organizational culture improves and the quality of self-evaluation is upgraded, independent evaluation should step aside from domains best handled by management and reorient its focus, e.g. it may move to a higher plane not yet served adequately by self-evaluation. In other words, "subsidiarity" - the organizing principle according to which matters ought to be handled at the lowest and least centralized competent level of authority – is an excellent principle to use in designing the relationship between independent and self-evaluation.

Self-Evaluation and Independent Evaluation Should Be Closely Connected

A combination of independent and self-evaluation encourages managers to design evaluable programs, i.e. adoption of clear goals, verifiable objectives and adequately funded monitoring and evaluation arrangements. For example, for results based management approaches to work evaluation should be built upfront into the design of interventions. This ought to be part and parcel of quality assurance for all public interventions.

Regular provision of independently validated information as to how the organization is using its resources; whether its interventions are addressing key priorities and whether they are delivering results implies that all major programs should be tracked and that accurate evaluation information should be published on regular basis so that stakeholders have access to clear, regular and credible information about organizational effectiveness.

At its best monitoring is geared to policy and program evaluation and re-appraisal. This means that monitoring is critical to management. But it also means that independent and self-evaluation add value to monitoring by applying criteria and standards to empirical findings; examining the merit and worth of results; explaining deviations from plans; taking account of changes in the context; reconsidering the assumptions on which the original intervention was based; ascertaining the continued justification of the intervention; guiding decision making regarding its adaptation to new circumstances; addressing unintended consequences; realigning roles and responsibilities of partners; etc.

Does Size Matter?

Less in evaluation can be more. Rapidly diminishing returns set in when independent evaluation adds to transaction costs without inducing much 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" (Arrow, 1974, pp.77-78).

Hence self-evaluation that serves and strengthens authority and reduces the incidence of unnecessary organizational errors should be privileged in resource allocation. On the other hand, the overall evaluation function combining independent and self-evaluation should not be starved of resources. Sufficient budgets should be allocated to allow a fair, valid, accurate and well documented assessment of overall organizational effectiveness.

To this end, monitoring and self-evaluation systems should be complemented by independent assessments that carry out spot checks and attest to the validity of self-evaluation claims. 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.

Conclusions

Independent evaluation helps to improve organizational performance. This is ultimately because organizational learning differs from individual learning. Especially in large, hierarchical organizations lessons drawn are not necessarily lessons learnt. Sound evaluation findings are not automatically adopted. Beyond knowledge acquisition the learning process requires attitudinal shifts and behavioral changes. For organizations this implies periodic changes in strategy, policies and procedures.

Inducing staff to observe currently agreed procedures and focus on corporate goals is required to achieve organizational effectiveness: this is the province of monitoring, auditing and single loop learning. But double loop learning is also needed, i.e. the strategies, policies, processes and procedures that shape organizational behavior require timely adjustment when the operating environment is turbulent or unstable: this is made much easier with the help of high quality independent evaluations that interrogate the continued validity of established strategic objectives and to reconsider the rationale of deeply ingrained business processes.

Conversely management systems that do not contribute to timely adaptation of strategic goals and operational guidelines are risky. Doing things right can destroy organizations that do not do the right things. Evaluation embedded in operational processes helps to ensure nimble adjustments of strategies and policies to changes in the external environment.

Self-evaluation is closer to the action so that it is more likely to impact on organizational behavior. However, it lacks the distance needed for objective assessment and it is likely to deteriorate in quality and it is vulnerable to capture by vested interests within the organization without an independent evaluation function tasked with attesting to the validity of its findings just as auditing does with respect to accounting.

In sum, evaluation independence defined in functional terms rather than simply in terms of evaluators' skills and dispositions is an essential ingredient of evaluation excellence. It helps to enhance the quality and credibility of evaluation products and it contributes to organizational transparency and accountability. The concept has been codified to ensure that evaluators are equipped to deliver high quality, uncompromising reports and to shield the function from capture and intimidation. All criteria of evaluation independence (structural, behavioral, protection from external influences and avoidance of conflict of interest) need to be put in place for the function to be genuinely independent.

However, independence is not isolation. Indeed, evaluation does not facilitate organizational learning if it fails to feed into strategic formulation, to amplify the voice of legitimate stakeholders or to provide credible and reliable performance information to the supreme governance authority. Independent evaluation also needs to leverage its impact through appropriate linkages to self-evaluation processes in order to generate "double loop learning".

This implies suitable linkages between independent and self-evaluation, including judicious protocols of professional interaction. It implies careful institutional design of organizational structures and business processes since without effective relationships among corporate oversight units charged with single loop learning and evaluation units tasked with double loop learning, incoherence and duplication may result. Finally, it goes without saying that 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.

Acknowledgements

This article draws on a paper commissioned and supported by the Independent Evaluation Office of the International Monetary Fund in 2012. The author thanks Matt Galen for helpful comments and generous editorial help.

Footnotes

[1] Utilization based evaluation approaches highlight the power of fully collaborative and participatory evaluation processes that builds trust and enlists the attention, knowledge and commitment of program managers and staff (Patton, 1997).

[2] The Glossary of Key Terms in Evaluation and Results Based Management issued by the Development Assistance Committee of the OECD 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."

References

Argyris, C. (1977). Double loop learning in organizations. *Harvard Business Review, 55(5)*, 115-125.

Arrow, K. J. (1974). *The limits of organization*. New York, NY: Norton.

Evaluation Cooperation Group (2010). Good practice standards on independence of international financial institutions' central evaluation departments. Retrieved from https://wpqr1.adb.org/LotusQuickr/ecg/Main.nsf/$defaultview/AA95B62CF943F82E4825774B003A82E0/$File/GPS%20on%20MDB%20Governance%20and%20Independence%20final%2007-02-10.pdf?OpenElement.

Hirschman, A. O. (1970). *Exit, voice, and loyalty: Responses to declines in firms, organizations, and states*. Cambridge, MA: Harvard University Press.

Love, A. J. (1991). *Internal evaluation: Building organizations from within*. Newbury Park, CA: SAGE.

Mayne, J. (2008). *Building an evaluative culture for effective evaluation and results management*. (ILAC Working Paper 8). Rome, Italy: Institutional Learning and Change Initiative. Retrieved from http://www.cgiar-ilac.org/files/publications/briefs/ILAC Brief 20 Evaluative Culture pdf.

Moore, M. H. (1995). *Creating public value: Strategic management in government*. Cambridge, MA: Harvard University Press.

Olson, M., Jr. (1971). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.

Perkins, D. N. (1993). Teaching for understanding. *American Educator, 17(3)*, 28-35.

Picciotto, R. (2005). The value of evaluation standards: A comparative assessment. *Journal of Multidisciplinary Evaluation, 2(3)*, 30-59.

Picciotto, R. (2011). The logic of evaluation professionalism. *Evaluation, 17(2)*, 165-180.

Pollitt, C. & Boukaert, Geert (2000). *Public management reform: A comparative analysis*. Oxford, England: Oxford University Press.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: SAGE.

Vedung, E. (1997). *Public policy and program evaluation*. New Brunswick, NJ: Transaction.

Weiss, C. H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation, 19(1)*, 21-33.

Appendix

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 • the Management, and/or • Board or • 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
	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 • staff, records, and project sites; • co-financiers and other partners, clients; and • 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: • 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; • 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.
		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 • hiring, firing, • term of office, • 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: • staff hiring, • promotion, pay increases, and • 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
	Preconceived ideas, prejudices or social/political biases that could affect evaluation findings	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; • 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

Note. Adapted from "Government Auditing Standards, Amendment 3" by U.S. General Accounting Office, 2002 "Glossary of Key Terms in Evaluation and Results Based Management" by Organisation for Economic Co-operation and Development, 2002; "Principles for Evaluation of Development Assistance" by Organisation for Economic Co-operation and Development, 1991; "Code of Ethics and Auditing Standards" by International Organisation of Supreme Audit Institutions, 2001; "Professional Practices Framework" by Institute of Internal Auditors, 2000; "The Conceptual Approach to Protecting Auditor Independence" by European Federation of Accountants, 2001; "Evaluation Guidelines" by Danish Ministry of Foreign Affairs, 1999; "Evaluation Guide" by Canadian International Development Agency, 2000.