CHAPTER 5

THE IEO AMONG ITS PEERS—LEARNING FROM COMPARATIVE EXPERIENCE

MIGUEL DE LAS CASAS AND ROXANA PEDRAGLIO

INTRODUCTION AND BACKGROUND

The Independent Evaluation Office (IEO) at age 20 is by far the smallest and youngest among sister evaluation offices in major international financial institutions (Table 5.1). Its evolution over the last two decades has been mostly inward looking, driven by its own experience and the changes undergone by the International Monetary Fund (IMF), and largely uninfluenced by developments in other evaluation offices and the evaluation community at large. Thus, this anniversary is a good opportunity to reflect not only on the IEO's own experience, but to seek to learn from the experience of larger, older siblings. While neither systematic nor exhaustive, this chapter analyzes the IEO's similarities and, more importantly, differences with evaluation offices in other international financial institutions. It also explores the possibility of the IEO incorporating strategic approaches or aspects of evaluation work that seem to have yielded good results elsewhere.

Admittedly, the IMF is unique in its functions. The IMF is not a development institution and, therefore, the nature of its relationship with member countries, its objectives, and its time horizons are quite different from those of other international financial institutions. This is reflected in the design, operations, and target audience of the IEO, which are in many ways distinct from those of peer evaluation functions discussed in this chapter. These dissimilarities must be taken into account when comparing the IEO with other evaluation offices. However, comparison is still a fruitful exercise as other institutions carry out activities that are comparable to the IMF´s surveillance and policy advice and all of these institutions provide capacity development and financial support to their member countries. Moreover, there is no doubt that the IEO grapples with many of the same dilemmas and problems faced by its peers in trying to use evaluation as a lever to trigger positive change in their institutions.

Another important preliminary consideration is that the mandate and design of each evaluation office is, of course, the result of the needs and preferences of each institution, of the circumstances at the time of their creation, and of their evolution over time. The IEO was created at a time when the IMF was a much less transparent institution, and many stakeholders were against the establishment of an independent evaluation function. In fact, the IEO was born only after years of discussions (IMF 1992). When the office was finally created, two preoccupations of IMF Executive Directors profoundly marked its design.[1] First, Executive Directors were concerned about the potential for the IEO's activities to interfere with the IMF's operations, particularly in the context of sensitive program negotiations. As noted in Chapter 2, this led to the introduction of the "non-interference clause" in the IEO's Terms of Reference (TOR) requiring that the IEO should "avoid interfering with operational activities including current programs."[2] This clause is unique among the evaluation offices discussed in this chapter and has a potentially key bearing on the scope of the IEO's activities. Second, IMF Executive Directors perceived evaluation offices elsewhere as too large, lacking independence, and excessively populated by staff from parent institutions. These perceptions resulted in the adoption of a business model for the IEO focused on ensuring independence, balanced composition of staff,[3] and a limited budget guided by a "less-is-more" approach.

This chapter draws on interviews with the Evaluation Department (EvD) of the European Bank for Reconstruction and Development, the Independent Development Evaluation (IDEV) of the African Development Bank, the Independent Evaluation Department (IED) of the Asian Development Bank, the Independent Evaluation Group (IEG) of the World Bank, the Independent Evaluation Office of the United Nations Development Program (IEO-UNDP), and the Office of Evaluation and Oversight (OVE) of the Inter-American Development Bank. In addition, the authors have analyzed evaluation reports from these evaluation offices with comparable themes and coverage, namely on the engagement

1 See IMF 2000a and IMF 2000b.

2 See IEO 2015, p. 2.

3 According to its TOR, a majority of the IEO's staff must come from outside the IMF. Relatedly, the IEO and all other comparators have adopted a model by which evaluations are led by staff of the evaluation offices. While all offices work with consultants, over the years the tendency has been to move away from outsourcing, relying instead on internal evaluators for the design and writing of core reports. This model is unanimously considered superior in terms of evaluations' accuracy and adaptability to the needs of the parent institution.

with fragile states and on the response to the global financial crisis,[4] as a basis for the discussion of differences and similarities with the IEO. The chapter also was informed by material compiled by the Evaluation Cooperation Group on evaluation functions across its membership.

IEO'S STRENGTHS

The "Less-Is-More" Approach

The IEO is a very small evaluation office relative to its peers by any metric, including human resources, budget, and, especially, output (Table 5.1). Compared to evaluation offices at other international financial institutions, it produces fewer reports each year, both in type and number.

TABLE 5.1. THE IEO AND ITS PEERS IN NUMBERS

	IEG (WB)	IED (ASDB)	OVE (IADB)	IDEV (AFDB)	EVD (EBRD)	IEO (UNDP)	IEO (IMF)
Reports per year	473	89	18	11	8–10	19	2–3
Budget (millions USD)	37.3	14.8	8.4	10.4	4.4	11.8	6.2
Budget ratio	1.4	1.9	1.4	1.8	0.8	0.2	0.5
Staff	109	60	25	40	18	32	15
Staff ratio	0.9	1.7	1.3	1.9	0.7	0.1	0.6
Date of establishment	1973	1978	1999	1980	1992	1967	2001

Sources: IEG, IED, OVE, IDEV, EvD, IEO-UNDP, IEO calculations.
Note: Budget (staff) ratio is defined as the budget (staff) of each evaluation office divided by the budget (staff) of its parent institution. FY2021 data.

This deliberate "less-is-more" approach has served the IEO well by boosting the depth and impact of its evaluation products, facilitating greater attention to each evaluation from the Executive Board, management, staff, country authorities, and other stakeholders. This approach relies on careful consideration of topic choice and timing, seeking to ensure that the workload generated by the presentation and

[4] The list of reports includes the work by Independent Evaluation Group (IEG) of the World Bank, the Independent Development Evaluation (IDEV) of the African Development Bank, and the Independent Evaluation Office of the United Nations Development Program (IEO-UNDP) on fragile states (IEG 2016; IEG 2014; IDEV 2020; IEO-UNDP 2013), the work by the IEG and the Independent Evaluation Department (IED) of the Asian Development Bank on the response to the global economic crisis (IEG 2012; IEG 2010; IED 2012: and IED 2011), and the Development Assistance Committee (DAC) guidelines for evaluation on conflict and fragility context (OECD 2012).

follow up of evaluations does not exceed the absorption capacity of active partners.[5] Overloading recipients would reduce the chances of reports catalyzing learning and change, as the Board's support for the evaluation function and management's and staff's attention to the design and execution of implementation plans would inevitably wane. It could be argued that the "disciplining" effect of potentially being subject to an evaluation is smaller when the total number of evaluations is smaller (although the scope of IEO evaluations has tended to be quite broad), but neither the IEO's "founding fathers" nor its subsequent directors were inclined to adopt a "more-is-more" strategy, for which a major overhaul of the IEO's structure and resources would be required.

In contrast, colleagues in other evaluation offices shared experiences of producing high, often excessive volumes of evaluation material, leading to somewhat mechanistic responses and limited impact on their parent institutions. In some cases, the issuance of numerous and excessively ambitious reports was seen as deriving from evaluation offices' need to justify their size and budget, as well as a tendency to want to "cover all the bases." In other cases, the large volume of evaluation material was demand driven, coming in response to requests from the Board, management, and staff. This may reflect governance structures in international organizations that make it easy to request additional streams of work, while refusing and prioritizing activities is more difficult. Colleagues in other offices have tried to palliate the problems associated with "overproduction" by offering easily absorbable, simplified reports, such as evaluation briefs and synthesis papers, by limiting reports' length, by having explanatory discussions with individual directors or subgroups of the Board, and by reducing the number of reports elaborated.

The types of evaluation reports produced by each office is determined, to a large extent, by the needs and peculiarities of each institution and the resources available to the evaluation function. Typologies are also dynamic, so it is difficult to compare over longer periods of time. The IEO's outputs[6] are closer to what other evaluation offices usually refer to as corporate and thematic evaluations, which could be defined as occasional reports focused on a particular activity, theme, aspect, policy, or process, and aimed at improving the performance of the parent institution. For example, the various reports on fragile states and the impact of the global financial

5 See IEO's "Possible Topics for Future Evaluations" document: https://ieo.IMF.org/en/our-work/work-program.

6 As discussed in Chapter 2, the IEO produces only three types of reports: full evaluations, short evaluations (recently created and still in pilot stage), and evaluation updates, which revisit issues approximately 10 years after an original evaluation.

crisis reviewed for this chapter were classified as major evaluations (IEG 2012; IEG 2010; IEG 2016; IEG 2014), corporate evaluations (IDEV 2020), thematic evaluations (IEO-UNDP 2013), corporate and thematic evaluations (IED 2011), and evaluation topical papers (IED 2012). In contrast to other offices, the IEO does not produce impact evaluations,[7] validations of self-evaluations carried out by IMF staff, periodic country or regional-level evaluations, systematic project evaluations, project completion reports, or sectoral evaluations. These evaluations consume a large portion of the resources at other offices and, to some extent, justify the difference in size. But even when comparing just evaluations of the type carried out by the IEO, the number of reports produced by other offices is much higher.

An additional advantage of the IEO's narrower approach and fewer product types is that it has more freedom to allocate resources to the most relevant topic at a particular time and to tailor the approach to the topic at hand. Some colleagues in peer offices expressed concern than an excessive variety of product types and rigidity in formats distracts from the core objectives of evaluation. At the same time, with high production volumes, having standardized typologies was considered to save time and effort by clarifying from the outset the parameters, procedures, and governance guiding each type of report.

Independence

The three external evaluations of the IEO (Lissakers and others 2006, Ocampo and others 2013, and Kaberuka and others 2018) found that the IEO has high levels of independence, both statutory and in practice, from the IMF's Executive Board, management, and staff. Evidence from interviews also suggests that the IEO typically operates with a relatively higher degree of independence than other evaluation offices. This is due in part to the preoccupation with independence at the time of IEO's creation, which was baked into the IEO's TOR, as discussed earlier. Several colleagues in peer offices saw themselves facing greater threats to their independence than the IEO, given the fact that their institutions are larger and more complex than the IMF and host a variety of pressure groups that try to influence their work.

The selection of evaluation topics is an area that illustrates the statutory independence of the IEO. For the IEO, as for all the evaluation offices included in this study,

[7] While the IEO looks into the outcome and broad impact of IMF activities, it does not conduct quantitative impact evaluations. This specific type of evaluation assesses how interventions affect outcomes, in an intended or unintended way, by analyzing counterfactuals. Impact evaluations are not always feasible and typically absorb a substantial amount of resources. See OECD's "Outline of Principles for Impact Evaluation": https://www.oecd.org/dac/evaluation/dcdndep/37671602.pdf.

the topic selection process involves informal consultations with Executive Directors and other relevant stakeholders on their views and preferences. Following these consultations, the IEO discusses topics with the Board's evaluation committee and publishes a list of around 10 possible topics for evaluation. The key difference is that the IEO Director is ultimately free to choose evaluation topics,[8] retaining autonomy to make a final decision up to the moment evaluations are launched. By contrast, in other evaluation offices, work plans require Board approval and, in some cases, Executive Directors—and even management—can request evaluations to be undertaken, perhaps reflecting differences in the intended target audience of evaluation reports. Another difference is that in other evaluation offices, topic selection is restricted by a more explicit set of criteria[9]—often contained in an evaluation policy—and by the need to produce recurrent periodic evaluation products. In addition to being more autonomous, the IEO's process seems to be quicker and more flexible. Colleagues in peer offices considered that their processes are too long, taking up to six months, and excessively rigid; in some cases, once the working program has been decided, the first two years of activity are basically fixed.

In general, interviewees saw their topic selection process as more rigid, more burdensome, and more exposed to internal and external influence than the IEO's. Some had initial lists of topics containing over 200 proposals. Others said that their topic selection process had become a consensus-building exercise aimed more at pleasing stakeholders than identifying the most crucial topics. A work program with a long time horizon is seen as favoring early planning and preparation of evaluations but also adding rigidity.[10] Some colleagues, however, did not view the need for Board approval of the work program (and the associated budget) as reducing independence. They thought it only logical for Executive Directors, as key recipients of evaluation reports, to choose evaluation topics that they consider useful for the institution, and they viewed final decisions on topic selection as the outcome of negotiations between their offices and the Board, rather than as imposed by the latter.

[8] The TOR provides the IEO Director with a clear mandate to choose topics and make decisions at arm's length from the Board and independent from management and staff. The IEO Director is chosen by the Board for a single non-renewable six-year term and cannot subsequently join the IMF staff. As is the case in other evaluation offices, the IEO Director can only be dismissed by the Executive Board, not by management.

[9] Examples of criteria used in peer offices, also considered by the IEO, are timeliness, relevance (materiality), strategic selectivity, evaluability, and usefulness.

[10] Some evaluation offices with comfortable budgets set aside a portion of their budget in order to be able to respond to last-minute requests for evaluation from their Boards. Other offices, however, cannot afford that luxury and react to contingencies (or Board requests) by rescheduling their working programs.

Similar to the topic selection process, some colleagues said they face more interference during the approach paper phase of their evaluations than is the case for the IEO. In some evaluation offices, early versions of approach papers are sent to and discussed with management before finalizing. While early consultation is intended to reinforce usefulness and receptiveness, it can also complicate the evaluation process. There have been cases of management interfering with the approach paper, even terminating an evaluation at this stage. Also, in some institutions, approach papers require Board approval and Executive Directors may insist on changes, although this practice seems to be uncommon. In the case of the IEO, approach papers (referred to as issues papers internally) are discussed with the Board only informally, providing Executive Directors with an opportunity to express their non-binding views. Similarly, the IEO seeks views on issues papers from management and staff. However, issues papers are not formally approved and Board agreement has been relevant only in cases where Executive Directors, management, or staff raised concerns regarding the non-interference clause.

Quality Control

To underpin the quality of its work, for each evaluation the IEO typically organizes three workshops, held at the beginning, midway, and close to the end of the evaluation process.[11] In these short events, a group of specialists discusses with the IEO the available evaluation material at each stage. The composition of these groups often varies from meeting to meeting and incorporates diverse perspectives, including from academics and other external experts, former IMF staff, former Executive Directors, and representatives from other international financial institutions and civil society organizations (CSOs). This process helps ensure that evaluations are based on analysis and evidence, that relevant information is not omitted, that alternative views are reflected, and that the content will be effective for external as well as internal audiences.

In contrast, peer offices use different mechanisms for quality control, in most cases, involving peer reviews and external advisory panels. Reviewers are internal—from the evaluation office, the parent institution, or both—and external—from academia,

[11] Beyond the quality control mechanisms described in this section, like other evaluation offices, the IEO shares with IMF staff a draft of all evaluation outputs for comments and factual corrections before they are sent to the Board. In this sense, evaluations go through the IMF´s well-structured review process, which constitutes an additional quality assurance. Moreover, the lower frequency and usually shorter length of IEO evaluations allow for greater scrutiny of their content. In all cases, the IEO retains ultimate responsibility for the findings and recommendations of the evaluation.

think tanks, and other institutions.[12] These reviews are quite formal, structured, and conducted at the start of the evaluation, often during the approach paper phase, and at the end of the process, when a draft report is ready. During the evaluation process there may also be contacts with reviewers, but those are informal and mostly dependent on personal relationships. In some cases, colleagues expressed concern that their validation processes have become very elaborate, time consuming, and bureaucratized; sometimes more focused on box-ticking than on reinforcing quality. Some considered their quality control mechanisms to be too internally oriented and lacking alternative views, while others saw their mechanisms as dominated by external experts who lack institutional knowledge.

The IEO seems to have found an equilibrium in its quality control, with a system that is less formal, but effective. While colleagues interviewed saw some advantages of peer reviews and advisory panels, they tended to prefer the IEO's model. They considered it less procedural and more apt to yield thoughtful, constructive discussions with a broader range of experts and stakeholders. They also appreciated the balance between external and internal knowledge and the relative speed of the process.

LESSONS TO BE LEARNED

Timing of Evaluations

The IEO has been freer than its peers when choosing evaluation topics, and when doing so, seeks to ensure that evaluations are timed to feed into rather than overlap with staff internal reviews. However, it has faced an additional constraint regarding the timing of its evaluations. As previously mentioned, when the IEO was created, the Board was preoccupied by the possible interference of the IEO's work on the operational activities of the IMF. This led to the adoption of a strict retrospective

[12] As an illustration, peer offices' quality control mechanisms for the reports consulted for this chapter were as follows: (i) the first evaluation by the World Bank's IEG on fragile and conflict-affected states (FCS) (IEG 2014) relied on an external advisory panel made up of two internal and two external reviewers, while the second (IEG 2016) had two peer reviewers, one internal and one external. IEG used three peer reviewers, internal and external, for both of its reports on the global economic crisis (IEG 2010 and IEG 2012); (ii) IDEV used four peer reviewers, three internal and one external, for its 2020 report on fragility and resilience (IDEV 2020); and (iii) IEO-UNDP relied on a combination of an internal reference group and an external advisory panel for its evaluation on FCS (IEO-UNDP 2013). The rest of the evaluations used similar mechanisms, but they were not specified in the reports.

approach to evaluation and to the introduction in the IEO's TOR of the so-called "non-interference clause," intended to preclude the evaluation of current programs.[13]

The IEO's retrospective approach and non-interference clause have limited its ability to engage in a timely way in some topics, mostly related to evaluation of multiyear IMF-supported programs. This has contributed to the IEO being slower than peers in extracting lessons from experience. For example, the IEG and IED reports on the global financial crisis reviewed for this chapter (IED 2011 and 2012; IEG 2010 and 2012), were conducted shortly after the crisis, in stark contrast with the timing of the IEO's evaluation, completed six years after the crisis (IEO 2014 and 2016).[14]

In the past, the non-interference clause also discouraged the IEO from undertaking early-stage evaluations,[15] which are issued by all its peers. Early-stage evaluations are conducted for recent interventions, allowing for a formative element.[16] These evaluations are typically well received, and indeed demanded, by the Boards and managements of other institutions. They provide extremely useful and timely inputs, often through the provision of interim reports to the Board. They can help guide operations, improve performance, reduce costs and risks, identify trends and opportunities, and lead to adjustments in the original assumptions underpinning the program or policy being evaluated. Most early-stage evaluations do not evaluate outcomes as they are conducted too early to be able to assess them and focus instead on processes and outputs at a certain time, assessing if the intervention was consistent with the underlying problem. Consequently, accountability is reduced, because the ultimate success of operations is not evaluated. However, as noted in Chapter 3, for Executive Directors, one advantage of early-stage evaluations is that they facilitate the Board's oversight function by narrowing the

---

[13] The clause currently reads: "In conducting its work, IEO should avoid interfering with operational activities, including current programs" (IMF 2015).

[14] The IEO published in 2011 an evaluation of the performance of IMF surveillance in the run-up to the global financial crisis (IEO 2011).

[15] The IEO has recently experimented for the first time with a hybrid model. The evaluation on the IMF's response to the COVID-19 pandemic (expected publication in the first quarter of 2023) was conceived as a full, retrospective evaluation but given the strong demand for early practical lessons, the IEO introduced a mid-point informal engagement with the Board, held in March 2022, providing an early opportunity to discuss emerging findings.

[16] Formative evaluations are conducted during the implementation phase of projects or programs and are intended to improve their performance, while summative evaluations are conducted at the end of an intervention (or a phase of that intervention) to determine the extent to which anticipated outcomes were produced (OECD 2002).

information asymmetry that exists in most institutions between Board members and management.

An objection sometimes raised against early-stage evaluations is the risk of prejudging the final conclusions of a subsequent, full, retrospective evaluation. However, colleagues with experience in other evaluation offices perceive this risk as limited since early-stage evaluations are, by definition, preliminary, explicitly subject to the ultimate evolution of many variables, and understood as such by their intended audience.

Even more ambitious in terms of providing quick guidance are real-time evaluations (RTEs).[17] These evaluations aim to provide feedback in real time to those responsible for implementation and they can be considered live learning. Crucially, they provide an opportunity for the adoption of early corrective measures in the midst of a crisis, when RTEs are normally most useful. However, they require a specific set of tools, capacities, and processes at both the evaluation office and the parent institution. Only two of the evaluation offices interviewed for this chapter (IDEV and IED) have experience with RTEs and they highlighted the difficulty and work-intensity of the task. Currently, the IEO and the IMF are not equipped to conduct RTEs or benefit from them, since the required real-time systems for reporting and early adoption of corrective action are not in place.

None of the IEO's peer offices have a legal framework with a provision like the IEO's non-interference clause, either in their TORs or evaluation policies, and none face any restriction regarding the timing of their evaluations, beyond the rigidity of the work programs previously discussed. Colleagues in these offices view the non-interference clause as constraining the IEO's independence and limiting the potential value added of its work, while recognizing the special macroeconomic sensitivity of the IMF's mission. The risk of interfering with the ongoing operations of the parent institution was not considered a problem—for some, it was even seen as a desirable outcome of evaluation work, useful to correct ongoing mistakes, thus demonstrating influence as opposed to interference. Many of the evaluators who were interviewed also said that a clause of this type would not be possible in their offices because they

17 RTEs are widely used in the field of humanitarian assistance where early adjustment of responses is crucial; see, for example, UNHCR (2002) and Cosgrave et al. (2009). RTEs are conducted as events unfold and are formative in nature. In contrast, early-stage evaluations are conducted retrospectively while providing early feedback and lessons, thus combining summative and formative elements. It is also important to differentiate between RTEs and monitoring. While the former are conducted by evaluation offices, monitoring is the responsibility of management and staff, although the evaluation offices in multilateral development banks often provide guidance on how to design monitoring processes that allow for subsequent evaluation work.

are constantly evaluating ongoing operations in member countries. Indeed, this is also true for the IEO's evaluations of the IMF's surveillance and capacity development activities, where evaluation work is less constrained, although still done with care not to comment on specific ongoing operations.

Thus, the IEO's non-interference clause is clearly an anomaly and each of the three external evaluations of the office have suggested it be reconsidered, relying on the IEO's strong track record and demonstrated judgement to avoid timing-related sensitivities.[18] In fact, when the IEO was created, these safeguards were understood as a temporary measure, allowing time for the IEO (then called EVO) to establish its credibility. At the time, when discussing Executive Directors' concerns about independence and interference, the evaluation group charged with designing the IEO said in its statement to the Executive Board: "Here, as the saying goes, 'the proof of the pudding will be in the eating.' Put more prosaically, the credibility of EVO has to be earned by its performance and the effectiveness of its safeguards will have to be demonstrated through its operations."[19]

Has the IEO earned its credibility? Looking back over two decades and scores of reports, it is reasonable to assert that the IEO has done so. Moreover, even if the non-interference clause were dropped altogether, the Board would still have the authority to stop the publication of an IEO report that it considered potentially disruptive. An alternative approach would be to clarify and more narrowly circumscribe the clause[20]—as was attempted in the past—for example, in the context of a new IEO evaluation policy, discussed later in this chapter. Such a document could clarify that "ongoing operations" refers to "ongoing programs," ensuring the clause is not an obstacle to the flow of information to the IEO and removing any questions about its ability to evaluate surveillance, capacity development, and early responses to crises like the COVID-19 pandemic.

[18] As an example, the IEO adopted the policy of avoiding single-country evaluations given the sensitivity of such evaluations.

[19] IMF 2000c; p. 1.

[20] The 2018 external evaluation of the IEO (Kaberuka and others 2018) recommended this approach: "The Board should review the IEO's Terms of Reference to a) […] and b) ensure that the scope of "operational activities, including current programs" does not restrict the IEO from conducting useful evaluations of ongoing activities of the Fund."

Methodologies and Tools

Compared to its peer offices, the IEO is not as systematic or rigorous in using evaluation methodologies. In fact, during interviews, one colleague expressed the view that "what the IEO does is research. Very interesting research and with an impact on the IMF, but research, not evaluation work."

The IEO uses some of the same bread-and-butter tools and sources of information employed by peer evaluation offices: literature reviews; semi-structured interviews with country authorities, Executive Directors, staff, management, and third parties; surveys aimed at wider ranges of participants and stakeholders; country case studies; and analytical, statistical, and econometric work. Text analysis, regularly used in peer offices for some time, has also been recently incorporated to the IEO's toolkit. However, other offices have a more complete set of evaluation methodologies at their disposal.

Colleagues in all evaluation offices consulted agreed that it is fundamental to underpin evaluations with a solid methodological base. Methodology, together with independence and perspective, is viewed as a key comparative advantage for evaluation offices when drawing lessons and making recommendations. Evaluation methodologies help provide an alternative approach that staff in parent institutions do not bring to the table. The use of appropriate methodologies also strengthens findings, eliminating biases and other flaws. Some offices much larger than the IEO count on methodological units that seek to ensure rigor in the use of these tools.

The IEO has relied heavily over the last two decades on the analysis of substantive issues, using economic and policy analysis tools and approaches similar to those of IMF staff. The IEO's work is typically more candid than the staff's reports, seeks to answer somewhat different questions, and reflects a wider range of views from outside the institution, but is not fundamentally different. This approach leaves largely unexploited the potential benefits of alternative evaluation methodologies.

At the same time, some peer evaluators warned against the risk of methodological dominance, as experienced in their evaluation offices. More specifically, some colleagues were concerned about the possible "fetishization of methodology" that can create an "illusion of rigor." This may turn evaluations into methodological box-ticking exercises, to the detriment of a deep understanding of the institutions' operations, leaving reports lacking in substantive content. Another potential problem with methodological dominance is a tendency to write for evaluators, with evaluations so methodology-driven and full of technicalities that they are unintelligible to nonexperts. Clearly, a balance is needed, so that evaluations can benefit from both a deep understanding of the challenges facing the parent institution and a sound evaluative methodology. In addition, the application and refinement of methodological tools needs to be seen as an ongoing effort in order for evaluation offices to maximize their value added and relevance.

As an example of methodological differences, the IEO's approach is unlike others with regard to the use of theories of change (ToC). Also called "program theory" and "results chain," a ToC is essentially a map or a log frame to illustrate how a certain intervention by the institution is supposed to lead to an expected outcome. ToCs are universally and systematically used by other evaluation offices.[21] The Organisation for Economic Co-operation and Development (OECD) recommends the use of ToCs to set out in clear cause-effect terms how interventions are intended to produce outputs, outcomes, and impacts, spelling out the conditions needed for success, and it has pointed to faulty ToCs as one reason why interventions fail to produce the intended results (OECD 2012).

Many colleagues in other evaluation offices considered ToCs as an integral element of their evaluations, although they are formalized to varying degrees and are not always published. For example, ToCs were used for the elaboration of all the evaluation reports reviewed for this chapter, although only one was published.[22] Interviewees regarded ToCs as an invaluable tool to understand how the policy or project to be evaluated is supposed to work, to guide and focus the evaluation process, and to strengthen the value added of the evaluation and facilitate implementation. At the same time, some colleagues believe that developing explicit ToCs can be time-consuming, potentially detract from thinking through the issues, and be misleading if misspecified, for example by putting too much focus on how activities were intended to work when first conceived rather than how they actually worked.

It can be argued that the IEO uses implicit ToCs in evaluating IMF operations and policies. These implicit ToCs would be predicated on the IEO's very deep understanding of the IMF's activities and the highly organized, structured, and relatively narrow nature of IMF operations. Nevertheless, as many colleagues in other evaluation offices assert, the process of making ToCs explicit—and more importantly,

[21] Within the IMF there is also some experience with this type of tool. For instance, a ToC (log frame) underlines the design of results-based management system used for monitoring and evaluating capacity development work. An overall ToC for the functioning of the IEO is provided in Chapter 4.

[22] IEG 2014.

discussing them with the staff and management of the parent institution—reveals missing links and causal relationships, allowing key evaluation issues and questions to emerge. In this sense, explicit ToCs are not a prerequisite, but their use substantially reduces the risk of important omissions and misinterpretations. ToCs can be construed as an explication of best practice, hence their use as evaluative framework. In addition, ToCs are not only valuable to strengthen the quality of evaluation reports, but their use contributes to the institutional learning sought by evaluation offices, as this type of deep understanding of causal chains is often taken for granted or overlooked by staff.[23]

Other methodological tools used by peer offices include formal impact evaluation techniques, such as randomized control trials, content analysis,[24] contribution analysis,[25] and process tracing.[26] Several offices also are studying how to incorporate different applications of recent advances in artificial intelligence and "big data." Regarding the acquisition of these techniques, experienced colleagues advocated for providing evaluators with training from outside their offices, promoting the

[23] Following the experience of discussing and learning about ToCs with their evaluation offices, the staff of several institutions have incorporated ToCs in their regular operations. Some evaluation offices even provide training to the staff and management of their institutions—and of course to their own evaluators—on ToCs, which in turn helps strengthen those institutions' evaluation culture.

[24] Content analysis is a research tool used to determine the presence of certain words, themes, or concepts within some given qualitative data, such as text. Using content analysis, researchers can quantify and analyze the presence, meanings, and relationships of certain words, themes, or concepts. As an example, researchers can evaluate language used within a news article to search for bias or partiality. Researchers can then make inferences about the messages within the texts, the writer(s), the audience, and even the culture and time surrounding the text. https://www.publichealth.columbia.edu/research/population-health-methods/content-analysis.

[25] Contribution analysis is an approach for assessing causal questions and inferring causality in real-life program evaluations. It offers a step-by-step approach designed to help managers, researchers, and policymakers arrive at conclusions about the contribution their program has made (or is currently making) to particular outcomes. The essential value of contribution analysis is that it offers an approach designed to reduce uncertainty about the contribution the intervention is making to the observed results through an increased understanding of why the observed results have occurred (or not) and the roles played by the intervention and other internal and external factors. https://www.betterevaluation.org/en/plan/approach/contribution_analysis.

[26] Process tracing in evaluation is an approach used to assess the impacts of selected interventions based on qualitative data. It focuses on demonstrating causal inference using in-depth analysis of an intervention. In process tracing, the evaluator's main objectives are to showcase evidence of the extent to which an intervention's key targeted outcomes have materialized and to investigate the causal mechanisms responsible for the outcomes. The main difference between process tracing and other theory-based evaluations is that in process tracing, the theory of change is much more explicit and detailed, and each hypothesized causal relationship is tested using empirical evidence. https://idev.afdb.org/sites/default/files/documents/files/Process%20Tracing%20as%20a%20methodology%20for%20evaluating%20small%20sample%20sizes.pdf.

integration of different perspectives that can later be adapted to the specifics of the institution. In their opinion, relying on “on-the-job learning,” while fast and convenient, may lead over time to a very closed and narrow approach to evaluation. Relatedly, several colleagues emphasized the importance of evaluation offices hiring staff with diverse professional profiles, so that evaluation products are enriched with alternative perspectives.

Finally, the COVID-19 pandemic impacted the operations of evaluation offices the same way it has affected the operations of parent institutions. While virtual engagement has increased access to individuals by making schedules more flexible and reducing the need for travel, peer evaluators opined that the quality of communication has been severely affected. They believed that the lack of face-to-face interactions hampered trust and understanding of the topics being evaluated. Overall, there was agreement that virtual engagement will be used much more frequently than in the past but will not replace field work. COVID-19 has also put pressure on international financial institutions to quickly provide information and assessments on how institutions are responding. Consequently, evaluation offices have adopted some of the alternatives discussed previously and incorporated new technologies, including new survey software, “big data,” and satellite imagery, to quickly generate and disseminate data and findings.

Interaction with Staff and Management

There is a risk that the high levels of independence of the IEO discussed above could come at the expense of a more productive cooperation with some stakeholders, which might impact the quality, usefulness, and impact of evaluation work. Other evaluation offices tend to have a more abundant and closer interaction with the staff and management of their parent institutions, in part due to the heavier relative weight these groups have as part of those offices’ target audiences. While the IEO also interacts in a similar way with staff and management, in other evaluation offices there seem to be more formal and systematic interactions at all stages of the evaluation process: topic selection, elaboration of the approach or issues paper, finalization of reports and implementation of recommendations, including the design of specific actions.

The IEO and most peer evaluation offices have seen an increased level of interaction with the staff of their parent institutions over time. As their reputation grew and their institutional roles consolidated, their work became more appreciated and the risk (actual or perceived) of loss of independence or capture by staff diminished. This was not a smooth and linear progression; like the IEO, other offices have experienced ups and downs in their rapport with staff, depending on the approach taken by those in charge of both sides of that relationship and the tension generated by various evaluations.

The quality of the relationship between an independent evaluation function and its parent institution is a shared responsibility, requiring reciprocal respect, good faith, and recognition of their respective mandates and institutional roles. Colleagues in other evaluation offices concurred that, to a large extent, it is up to evaluators, through their policies and professionalism, to nurture a cordial and productive relationship with the relevant interlocutors, strengthening the focus, utility, credibility, and buy-in of the findings and recommendations while safeguarding the integrity of the evaluation function. In this regard, each director or chief evaluator, in the IEO and in peer offices, has had a particular preference regarding the appropriate level of in-practice independence, using the wiggle room provided by the formal independence framework.

The key is to find a balance between desirable consultation and interference. A discontinuous or unduly confrontational relationship with staff and management—which, at some point, has been a problem faced by all evaluation offices—may impede information flows and hamper the evaluation function. At the other end of the spectrum, experience shows that excessive interplay with the staff of the parent institution may erode the evaluation office's capacity to provide candid advice, out of concern to maintain the relationship or because the evaluation office played a role in the design of an intervention and is thus a biased observer rather than an impartial one.[27] As a result, in evaluation reports, messages extracted from evidence may be weakened and recommendations diluted. In general, there was also agreement among colleagues that the risk of compromising independence is compensated by the advantages of maintaining a constructive and sustained relationship—provided that the evaluation office is always careful to draw a line and push back if the line is being crossed.

[27] Unlike the IEO, some offices get involved in the operations of their parent institutions from the start, sharing lessons from previous evaluations, which can help improve operation design, and ensuring that the data and structure needed for later evaluation will be generated.

To clarify the relationship with staff and other aspects of their operations, peer evaluation offices have relied on explicit evaluation policies.[28] These Board-approved documents typically contain a unified set of rules of engagement with the parent institution, developing the norms included in other documents such as TORs, or complementing them, or both. They also contain evaluation guidelines, principles, standards, objectives, criteria, responsibilities, processes, and offices' mandates, which guide and elucidate evaluation work. In interviews, colleagues in other offices said they considered evaluation policies to be fundamental; this was especially true for those who survived for years with their TORs as their only legal anchor.

The IEO has a range of documents that cover various aspects of what would typically be included in an evaluation policy. These include, for example, the TOR for the office and separately, for the Director, terms and conditions of employment for IEO staff, and guidance for IMF staff on how to cooperate with the IEO.[29] However, adoption of a more comprehensive Board-approved policy document would gather this disparate material in a single articulated document, fill in gaps, or elaborate on areas of ambiguity. An evaluation policy could, among other things, underpin the current guidance to IMF staff on how to relate to the IEO and help delineate the applicability of the non-interference clause.

Closely tied to the type of relationship the evaluators have with staff are the goals pursued by the independent evaluation function. In this regard, all offices agreed that the main objective of evaluation should be promoting and facilitating constructive change in the evaluated institution, emphasizing learning more than accountability, although both goals are important and mutually reinforcing. Albeit with some nuances, all evaluation offices concurred with the idea that the primary goal of evaluation is learning—understood as the extraction and absorption of implementable lessons from a robust analysis of past experiences. Some colleagues saw accountability as forced knowledge, while evaluation offices typically aim at a more persuasive approach to learning. A hard accountability approach ("naming and shaming") has proven less fruitful and constructive as it focuses too much on individual performance rather than outcomes and inevitably hampers cooperation,

---

[28] All comparator evaluation offices have some form of evaluation policy posted on their websites, although these documents take different forms and names: evaluation principles, evaluation framework, evaluation brochure, or evaluation policy framework. IDEV used to have an additional document, referred to as evaluation strategy, bridging between the high-level, permanent evaluation policy and the evolving work program, but this practice was discontinued. Some UN agencies (UN Women, UNHCR, UNFPA) and national departments continue to utilize evaluation strategies. The list of references includes links to peer offices' evaluation policies.

[29] All these documents are available at the IEO website: https://ieo.IMF.org.

communication, and information flows. The IEO, like other evaluation offices, has gone through different periods in which the balance between accountability and learning shifted back and forth, in part reflecting the personality and attitude of the stakeholders involved, including the Board, the IEO, and the IMF's management and senior staff. However, from its inception, learning has been at the core of the IEO's mandate.[30]

Dissemination

At the most basic level, dissemination starts with a concise, well-written report. Typically, IEO reports are comparable to those of its peers in terms of length and readability. However, the IEO is not immune to a tendency towards long and very detailed evaluation reports, a point made repeatedly in external evaluations of the IEO and in management and staff comments on evaluation drafts. For external dissemination, as the experience of colleagues shows, communication can be more effective and efficient with greater use of visual tools and infographics and audience-tailored synopses centered on main messages.

It is mainly in the area of engaging in discussion and explanation where there seems to be room for the IEO to learn from colleagues. While the IEO has substantially increased its outreach initiatives over the years, it devotes much less time and effort than other offices to "in-reach," understood as internal (within the IMF) dissemination of its findings and proposals. In other institutions, once evaluations have been considered by the Board, communication and dissemination plans are put together for each report and discussions with staff are organized on the topics evaluated, sometimes going beyond the reports and their recommendations. Thus, evaluations are used to trigger a conversation on the substantive issues, facilitating learning and increasing the traction of reports.[31]

In-reach is beneficial in at least three ways. First, for each individual evaluation, it promotes learning by raising the awareness of issues among staff, providing a channel to explain the rationale behind recommendations, and thus raises the chances of successful implementation. Second, for the evaluation office,

---

[30] "We are cognizant of the need to create an environment within which EVO can contribute to a "learning culture" within the Fund, where Fund staff and management are as receptive as is possible to the observations and recommendations of independent evaluation. This is crucial if the recommendations arising from EVO evaluations are to be integrated into the work of the Fund in a timely manner, which is the ultimate objective of this exercise" (IMF 2000c; p. 1).

[31] Some peer offices, like IED and IDEV, organize periodic events (for example, evaluation weeks), which contribute to the dissemination and better understanding of their reports and of evaluation work in general.

dissemination helps raise the profile and appreciation of the evaluation function,[32] improving collaboration and rapport with staff, again boosting the likelihood of meaningful change in the future. Some colleagues asserted that a robust, high-quality evaluation, if poorly disseminated, may have little value or impact, as its chances of triggering change fall dramatically. Finally, for evaluators, dissemination is also beneficial as an opportunity to witness the impact of their work and reinforce their visibility and presence.

The IEO could improve its effectiveness by further strengthening the dialogue with staff and possibly Executive Directors on the main elements of those evaluations already discussed by the Board. This would contribute to increased understanding and ownership of the rationale behind recommendations and, thus, to a more effective design and execution of Management Implementation Plans (MIPs). Peer offices have achieved this through the organization of seminars and workshops with relevant interlocutors in their parent institutions.[33]

Beyond in-reach and outreach activities for individual reports, other evaluation offices devote more time and resources to the analysis, synthesis, and dissemination of the stock of knowledge and lessons from evaluation. IEO has now started to do this (see Chapter 2), but in a less systematic way than its peers. These products are shorter, easily digestible, and aim to be readily usable by the Board and other stakeholders in facing ongoing issues. As practiced by other offices, they typically require little or no new evidence gathering, as they rely on previous evaluations and are often supported by machine learning techniques that facilitate the analysis of large quantities of evaluative material. They may be demand or supply driven, and, in some institutions, they can take the form of a statement from the head of the evaluation office for the relevant Board discussion.[34] For example, IED and IEO-UNDP have synthesis and lessons series, IEG produces synthesis papers, EvD has a "Connecting the Dots" series, and IDEV publishes evaluation briefs and highlights and has a lessons-learned series.

[32] Some offices, in addition to accountability and learning, are mandated to strengthen evaluation culture in their parent institutions (and member countries) by providing training, but this is not the case for the IEO.

[33] The IEO has developed a successful series of seminars on economic issues, typically attended by IMF staff, personnel of the offices of Executive Directors, country authorities, and external participants. These seminars could provide a suitable forum for these discussions.

[34] In some institutions, such as IEG, the Director of the evaluation office can, motu proprio, issue a statement for any discussion of the Board.

Presentation of Reports and Follow-up

All peer offices submit their evaluation reports—including findings and recommendations—to their respective Boards or relevant committees. Executive Directors decide whether to endorse recommendations, normally after the institution's management team has expressed a more or less explicit opinion on each one of them. The IEO is one of only two institutions[35] where every evaluation report is discussed in full Board meetings—in the case of the IEO, presided by the Managing Director or a member of the senior management team—while all other evaluation offices present their work to Board committees that oversee evaluation and institutional effectiveness. The IEO's more intense Board engagement is possible thanks to the "less-is-more" approach described previously. Views of colleagues in offices that present their work to Board committees were mixed. Some appreciated the advantages of the IEO-IMF approach, including that meetings are chaired by management and followed by a formal summing up. Others did not see it as an important advantage, asserting that committees devoted more time and effort to evaluation reports than the full Board. Besides, they noted that in most cases committees have the same composition as the Board and, if necessary, once committees have discussed a report, it could be elevated to the full Board.

Over time, the IMF and the IEO have developed a robust system for monitoring the implementation of recommendations endorsed by the Board.[36] While all peer evaluation offices have an established procedure by which management is responsible for elaborating and executing an implementation plan, there are some important differences:

- In the case of the IEO, management has six months, from the date of the Board discussion, to present the corresponding MIP to the Board, translating Board-endorsed recommendations into concrete actions with a timeline and an accountability framework. In some other offices, this period is shorter, or even nonexistent, as implementation plans are presented to the Board or the relevant committee together with the evaluation report. This is only possible because, under this model, evaluations are discussed in depth with management before Executive Directors receive the reports, and management is given a very short period of time to come up with an implementation plan. In general, evaluators

[35] The other one is IEO-UNDP. However, comparison is difficult as the UNDP's Executive Board is nonresident and usually meets three times a year.

[36] For further details on the IEO's follow-up framework, see Chapter 4.

preferred a process like the one followed by the IEO, in which management has a relatively long period of time after the Board meeting to consider the Board's reaction to the recommendations and carefully develop and build consensus around the implementation plan. Accelerated processes have led in the past to excessively defensive reactions by management teams and suboptimal implementation plans that do not adequately respond to recommendations and guidance from the Board. Besides, given the Board's workload, management teams—the evaluated—tend to dominate the follow-up process, as the Board cannot monitor in detail the interpretation and subsequent implementation of recommendations in all evaluation reports.

- The IEO's system requires the Board's approval of the proposed MIP, which is not the case in other international financial institutions. Board involvement at this stage incentivizes more ambitious responses by management and a more dedicated subsequent implementation by staff.
- Unlike other institutions, at the IMF, periodic reports on implementation of recommendations—the annual Periodic Monitoring Reviews (PMR)—are also discussed by the full Board, not just the evaluation committee, boosting the Board's opportunity to raise concerns when implementation is offtrack, and to question staff's commitment to the process.

The IEO's system is also an outlier from other institutions in that responsibility for monitoring implementation lies with a third party, the Office of Internal Audit (OIA). In other institutions, evaluation offices themselves are charged with—or at least heavily involved in—the monitoring process. Having a third party oversee the monitoring of the implementation of recommendations brings objectivity and neutrality, and OIA is generally seen as fulfilling this role carefully and effectively. Nevertheless, most colleagues considered it preferable to keep that function within the evaluation office. In their view, evaluators are better positioned to judge whether implemented actions follow the spirit and substance of the original recommendations, or if the implementation process has become a perfunctory exercise. Another positive aspect of the model followed by other offices is that it promotes continuing contacts between evaluation offices and management and staff on how to better advance implementation, giving these offices an opportunity to contribute and keep implementation in line with the intended objectives.

CONCLUSIONS

Despite fundamental differences in the operations and mandate of their parent institutions, in their size and budget, and in their respective histories and responsibilities, a comparison of the IEO with peer evaluation functions is revealing and rich in lessons. Relative to comparable offices, the IEO is young and small, although it has achieved a remarkably large and positive impact in the IMF, as discussed in Chapters 2 and 4. It enjoys high levels of independence, has developed efficient mechanisms for quality control, and effectively monitors implementation of Board-endorsed recommendations.

At the same time, learning from the experience of older siblings, the IEO could strengthen its impact and reputation in a variety of ways. Consideration might be given to the following:

- The retrospective approach to evaluation and the non-interference clause have limited the IEO's ability to evaluate certain topics in a timely manner, thus reducing the value and applicability of findings and recommendations. While recognizing the unique sensitivity of the IMF's work, these restrictions could be relaxed or clarified, or both, to maximize the potential for learning through evaluation of ongoing activities.

- Incorporating formal evaluation methodologies like theories of change and impact evaluation techniques would strengthen the richness and robustness of IEO's reports and could be done in a cost-effective and tailored manner. Staff training with outside experts and a closer relationship with peer evaluation offices would help in this respect.

- Compared with peers, the IEO appears to have a more distant relationship with staff of its parent organization. Building on its own and others' experience, it is incumbent upon the IEO and IMF staff and management to continue to develop a collaborative, mutually respectful, and faithful relationship, with the aim of facilitating the ultimate objective of IEO evaluations: making the IMF more effective. Nevertheless, there may be room to clarify the principles guiding this relationship.

- Greater emphasis on dissemination, discussion, and explanation of its findings and recommendations to external and internal audiences, especially IMF staff, would increase the traction and appreciation of the IEO's evaluation work. Moreover, the IEO could contribute further to ongoing policy discussions, not only on policy documents prepared in response to its own recommendations but more broadly, by providing timely and relevant lessons distilled from previous evaluation evidence.

- While the IMF and the IEO have developed over time a robust system for monitoring the implementation of Board-endorsed recommendations, giving the IEO a greater role in this process, either in full or jointly with the OIA, may further strengthen this system.

- An evaluation policy like those of peer offices would be a useful addition to the IEO´s framework. This unified and coherent document would guide evaluation work, fill in gaps, clarify aspects such as the non-interference clause, evaluation criteria, and the types and timing of evaluations, while also promoting better understanding and better relations between the IEO and the IMF staff.

It should be recognized that additional resources would be needed for some of these initiatives, particularly to add evaluation methodologies, increase dissemination, and increase the IEO's role in monitoring. There is little scope to redeploy existing resources from a very lean operation that has not expanded since its inception 20 years ago.

The IEO´s 20th anniversary is an opportunity to address some of the legacies from the time of its establishment, recognizing the strengthening of the IEO's credibility and reputation and the IMF's progress in embracing transparency and a learning culture. Beyond suggestions for future improvement, this chapter highlights the importance for the IEO to proactively seek communication and shared learning with peer offices, and the evaluation community more broadly, at least over the next 20 years.

REFERENCES

Cosgrave, John, Ben Ramalingam, and Tony Beck, 2009, "Real-time evaluations of humanitarian action. An ALNAP Guide. Pilot Version," Overseas Development Institute.

Independent Development Evaluation of the African Development Bank Group (IDEV), 2020, "Evaluation of the AfDB's Strategy for Addressing Fragility and Building Resilience in Africa (2014–2019)," June (Abidjan, Côte d'Ivoire).

__________ (at the time Operations Evaluation Department), 2012, "Evaluation of the Assistance of the African Development Bank to Fragile States."

Independent Evaluation Department of the Asian Development Bank, 2012, "The Implications of a Global Financial Crisis for Asia and the ADB: Lessons from Evaluation," June.

__________, 2011, "Special Evaluation Study on Real-Time Evaluation of Asian Development Bank's Response to the Global Economic Crisis 2008–2009," June.

Independent Evaluation Group of the World Bank Group (IEG), 2010, "The World Bank Group's Response to the Global Economic Crisis—Phase 1," (Washington: World Bank).

__________, 2012, "The World Bank Group's Response to the Global Economic Crisis—Phase 2," (Washington: World Bank).

__________, 2014, "World Bank Group Assistance to Low-Income Fragile and Conflict-Affected States," (Washington: World Bank).

__________, 2016, "World Bank Group Engagement in Situations of Fragility, Conflict, and Violence," (Washington: World Bank).

Independent Evaluation Office of the International Monetary Fund (IEO), 2011, *IMF Performance in the Run-Up to the Financial and Economic Crisis: IMF Surveillance in 2004–07*, (Washington: IMF).

__________, 2014, *IMF Response to the Economic and Financial Crisis*, (Washington: IMF).

__________, 2016, *The IMF and Crises in Greece, Ireland, and Portugal*, (Washington: IMF).

__________, 2018, *The IMF and Fragile States*, (Washington: IMF).

Independent Evaluation Office of UNDP, 2022, "Annual Report on Evaluation 2021," June (New York: UNDP).

International Monetary Fund, 1992, "Establishing an Evaluation Office in the Fund" EBAP/92/166 (Washington: IMF).

__________, 2000a, "Making the Fund's Independent Evaluation Office Operational" EBAP/00/84 (Washington: IMF).

__________, 2000b, "Making Fund's Independent Evaluation Office Operational" EBM/00/81 (Washington: IMF).

__________, 2000c, "Statement by the Evaluation Group on Making Fund's Independent Evaluation Office Operational" Gray/00/686 (Washington: IMF).

__________, 2015, Terms of Reference for the Independent Evaluation Office of the International Monetary Fund, revised, September. (Washington: IMF). https://ieo.IMF.org/-/media/IEO/Files/terms-of-ref/tor.ashx.

Kaberuka, D., Chia Der Jiun, and Pernilla Meyersson, 2018, "Time for a reboot at a critical time for multilateralism—The Third External Evaluation of the IEO," July (Washington: IMF). https://ieo.IMF.org/en/our-mandate/External-Evaluations.

Lissakers, K., Ishrat Husain, and Ngaire Woods, 2006, "Report of the External Evaluation of the Independent Evaluation Office," March (Washington: IMF). https://ieo.IMF.org/en/our-mandate/External-Evaluations.

Ocampo, J.A., Stephen Pickford, and Cyrus Rustomjee, 2013, "External Evaluation of the Independent Evaluation Office," January (Washington: IMF). https://ieo.IMF.org/en/our-mandate/External-Evaluations.

Organisation for Economic Co-operation and Development (OECD), 2002, "Glossary of Key Terms in Evaluation and Results Based Management."

OECD, 2012, "Evaluating Peacebuilding Activities in Settings of Conflict and Fragility: Improving Learning for Results," DAC Guidelines and References Series, (OECD Publishing).

United Nations Development Program Evaluation Office, 2013, "Evaluation of UNDP Support to Conflict-Affected Countries in the Context of UN Peace Operations," January.

United Nations High Commissioner for Refugees, 2002, "Real-time humanitarian evaluations" EPAU/2002/05.

Evaluation Policies of IEO Peer Offices:

- IEG-WB: Evaluation Principles: https://ieg.worldbankgroup.org/sites/default/files/Data/reports/WorldBankEvaluationPrinciples.pdf and Evaluation Framework: https://ieg.worldbankgroup.org/methodology.
- IED-AsDB: Evaluation Brochure: https://www.adb.org/publications/independent-evaluation-accountability-and-learning-lessons-development-effectiveness.
- OVE-IDB: Evaluation Policy Framework: https://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=EZSHARE-872199154-11142.
- IDEV-AfDB: Evaluation Brochure: https://idev.afdb.org/en/page/related-page/about-us.
- IEO-UNDP: Evaluation Policy: http://web.undp.org/evaluation/policy.shtml.