

ANNEX 2

Evaluation Questions, Methods, and Sources of Evidence

This IEO evaluation set out to assess the relevance, quality, and utilization of self-evaluation in the IMF. The focus of utilization was on learning and improved institutional effectiveness, and in some cases also on accountability.

Evaluation Questions

- What provisions exist for self-evaluation in the IMF? Are assessments done in a systematic way in terms of methods and periodicity? How are findings documented? Are systems in place to draw conclusions and lessons that can be used to improve future work? How do the provisions in the IMF compare to those in other IFIs?
- What are the provisions for self-evaluation of the IMF's core functional areas (surveillance, lending, and capacity building)? What has been the experience to date with self-evaluation in each of these areas?
- Are self-assessments undertaken within each department? At the departmental level, are assessments done in a systematic way in terms of methods and periodicity? How are findings documented? Are there systems to draw conclusions and lessons that can be used to improve future work? What mechanisms exist to incorporate these lessons into operational work? How is this monitored? Are informal systems in place that play a similar role to formal self-evaluation processes?
- With respect to specific IMF supported-programs, what has been the experience with EPAs and EPEs? Who conducts these evaluations? How are objectivity and quality being assured? How well do the evaluations cover different aspects of program engagement? Do the evaluations distill meaningful lessons for IMF operations? What systems are in

place to ensure that lessons are being incorporated in future programs?

- To what extent and in what way do thematic reviews engage in self-evaluation of IMF strategies, policies, or practices? What systems are in place to disseminate and follow up on lessons distilled through these exercises?
- What types of self-evaluation is undertaken for the IMF's capacity-building activities, specifically technical assistance (TA) and external training? What systems are there for dissemination and follow up of lessons?

Evaluation Methods

The IMF does not have an institution-wide policy for self-evaluation or a formal, institution-wide self-evaluation system. Thus, the IEO defined a paradigm against which to assess each product or process evaluated with respect to coverage, quality, and utilization. Also, the IEO assessed compliance in those cases where there were specific guidelines (such as for EPAs and EPEs).

Evaluation of EPAs and EPEs

The IEO undertook a desk review of the 49 country self-evaluations undertaken between August 2005 and December 2013, including 3 Joint EPA/EPE reports and 7 EPA updates (see [Table 2](#) in the main report). Thirty of these reports covered programs primarily under concessional facilities (the Poverty Reduction and Growth Facility and the Extended Credit Facility) and 19 covered programs under the General Resources Account.

Each self-evaluation report was rated on the quality of the assessment according to a series of criteria covering the following broad headings: (i) the rationale for IMF program involvement; (ii) program design;

(iii) effectiveness of IMF involvement, including the appropriateness of conditionality; (iv) forward-looking strategy; and (v) overall judgment on the report, including the extent to which it identified clear lessons (see [Annex 1](#) of Goldsbrough, 2015 for the full set of assessment criteria). A four-point scale (“not discussed,” “weak,” “adequate,” and “very good”) was used to rate how well the reports performed under each criterion. In order to maximize the scope for comparisons, the criteria were designed so that most were applicable to both EPAs and EPEs, but some were applicable to only one type of report, reflecting specific mandates in their respective guidelines. The extent of any disagreements with the authorities or within the IMF Board (based on a review of the contents of the Summing Up) was also considered. The desk review could not assess whether particular judgments made in the reports were correct—since this would require a separate country evaluation—but focused on how well particular issues were explored, the quality of the supporting analysis, and the clarity of the lessons drawn.

Structured interviews were also conducted with senior Staff in area and functional departments involved with the EPA and EPE exercises as well as with a number of evaluation team leaders. The interviews were based on a standard set of questions sent in advance that covered (i) the usefulness of the EPA and EPE exercises as tools for internal self-evaluation and learning as well as areas for improvement; (ii) appropriateness of the country coverage; (iii) independence of the evaluation teams; and (iv) arrangements for following up on lessons generated by the evaluations.

Evaluation of policy and Thematic Reviews

The IEO undertook a desk study of 23 policy reviews and 35 thematic reviews that were conducted during 2006–13. The sample comprises about 110 documents, including background papers. Policy reviews are those so identified by the IMF in Table 4 of IMF (2015); however, the desk study excluded reviews in this list that clearly addressed technical issues of IMF financial policy and thus would not be expected to include self-evaluation—such as general reviews of quota and annual reviews of the IMF’s strategy on overdue financial obligations. The study is discussed in detail in Stedman (2015), and the full list of documents in the sample is provided in [Annex 2](#) of that paper.

First, the objectives and terms of reference of each review were assessed to determine whether self-evaluation

was undertaken by design, as an explicit part of the review. Those that set out to assess, evaluate, or review the effectiveness or adequacy of an IMF policy, instrument, or activity were considered to incorporate self-evaluation as an explicit part of their purpose.

Each review was then assessed to determine whether it included each of the following aspects of self-evaluation: (i) presented evidence about past experience, for example, in terms of compliance with policies or the quality of activities or products; (ii) conducted analysis in order to reach conclusions about the effectiveness of policies, practices, inputs, outputs, or outcomes; (iii) drew lessons from this analysis; and (iv) incorporated these lessons in its proposals for future policies or operations. To assess utilization, consideration was given to whether lessons led to policy or operational reforms. To understand the focus of self-evaluative work, each review was examined to assess the extent to which it addressed institutional issues, such as the effectiveness of a facility or the consistency of program conditions, as opposed to Staff practices in executing the institution’s work, such as how Staff engaged with authorities.

Evaluating Self-Evaluation of Capacity Development

The IEO examined the self-evaluation of IMF capacity-development activities in the period 2006–13. It assessed the mechanisms in place for self-evaluation of technical assistance (TA) and training to determine the degree to which assessments addressed the relevance, effectiveness, and impact of these activities.¹

- Relevance refers to the extent to which the TA/training is suited to the priorities and policies of the recipient countries and the IMF.
- Effectiveness refers to the quality of the TA/training output (i.e., whether the TA/training is designed well and delivered effectively, including through interaction and/or collaboration with country authorities and donors) as well as the TA/training outcome (i.e., the extent to which the TA/training improved the knowledge of the recipients).

¹This approach is based on the standard principles for evaluating development assistance developed by the Development Assistance Committee of the Organisation for Economic Co-operation and Development (OECD-DAC); see <http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>. The OECD-DAC principles include two other criteria—efficiency and sustainability—that are less directly relevant to the IMF.

- Impact refers to changes or effects resulting from the TA/training. Impact usually unfolds in stages, with the first stage being the extent to which the TA/training recipients make use of their increased knowledge and the final stage being the impact on macroeconomic performance resulting from recipients' use of this knowledge.

In September 2014, the Board approved a statement of policies and practices with regard to the delivery of capacity development activities, including monitoring and evaluation. This statement could be considered as a benchmark for assessing the self-evaluation of capacity development going forward.

Sources of Evidence

The IEO gathered information through reviews of documents from the IMF and other international organizations; and past IEO evaluations; interviews; and

surveys. It conducted semi-structured interviews with senior IMF officials in all area and functional departments and with a random sample of 44 IMF staff in grades A13–A15; and it interviewed Executive Directors from 21 constituencies. The evaluation team also conducted a survey of IMF staff.² Information on other international financial institutions was gathered by reviewing publicly available information as well as by surveying and interviewing senior Staff of the evaluation offices in these organizations. Background Papers were prepared on three topics: assessments of IMF-supported programs; self-evaluation of policies and other institution-wide issues; and the procedures in place for self-evaluation of capacity building activities. The team also consulted with evaluation experts, officials from member countries, and former senior IMF staff through interviews, workshops, and seminars.

²The results of IEO survey of IMF staff are summarized in Background Document 1, which also includes sample questionnaires used for interviews.