

Self-Evaluation of IMF Policies and Other Institution-Wide Issues

38. This section assesses self-evaluation of policies and other institution-wide issues at the IMF.²³ There were no explicit requirements to conduct free-standing self-evaluation of policies or institution-wide issues. However, self-evaluation was conducted as an input or as a by-product of the periodic policy reviews and other thematic reviews that the IMF prepared as the basis for forward-looking policy development and adjustments to operations.²⁴ This section examines the extent to which these policy and thematic reviews incorporated evaluative analysis, as well as their quality and utilization.

39. Policy and thematic reviews conducted over the evaluation period spanned the IMF's main activity areas and key operational and policy issues. On *lending*, the topics covered by policy reviews included conditionality, specific facilities or instruments (the Flexible Credit Line, facilities for low-income countries, and the Policy Support Instrument), access policy, transparency, and debt limits. Thematic reviews addressed issues including the IMF's financing role (2008), crisis programs (2009), and monetary policy in a subset of member countries (2013). On *surveillance*, policy reviews examined, inter alia, surveillance, the Financial Sector Assessment Program, data provision, standards and codes, and

data standards. Thematic reviews included an analysis of initial lessons from the crisis (2009), examination of IMF involvement in the G20 Mutual Assessment Process (2011), and consideration of macroeconomic issues in fragile states (2011) and small states (2013). The IMF also conducted thematic reviews of *technical assistance*, as discussed in Chapter 5 below.

Self-Evaluation Is a Component of Many Policy and Thematic Reviews

40. The IEO examined 58 policy and thematic reviews prepared during 2006–13 to assess the extent to which they incorporated self-evaluation as means of learning, and enhancing transparency and operational effectiveness. This analysis, summarized in Table 2, addressed six questions for each review. First, the IEO asked whether self-evaluation was an explicit goal of the review. For each review, whether or not it had an explicit self-evaluation goal, the IEO then examined whether it presented evidence about past experience; analyzed the effectiveness of policies or practices; and drew lessons from this analysis to guide future policies or operations. To assess utilization, the IEO explored whether these lessons then fed into policy or operational reforms. Finally, it asked whether the evaluative work addressed Staff practices, in addition to institutional or policy issues.

41. Nearly 60 percent of the policy reviews (but only 17 percent of thematic reviews) explicitly included self-evaluative aims in their objectives or terms of reference; in each of these reviews, self-evaluation was at least a building-block of a broader discussion.²⁵ For

²³This section draws on Stedman (2015), which examined 23 policy reviews and 35 other thematic reviews (comprising 110 documents, including background papers), and where the documents are listed.

²⁴The analysis in this section covers two categories of reports: a set of “policy reviews” that were mandated by the Board to take place on a periodic or as-needed basis, as listed in IMF (2015), Table 4, and other reviews that took place on an ad hoc basis, referred to here as “thematic reviews.” In April 2015, in the context of the FY2016 budget, the Board decided to consolidate reviews and shift most of them to an as-needed basis with a periodicity of five years or more. Reviews of surveillance, conditionality, and capacity building will take place on a five-year schedule; review of the Debt Sustainability Framework will continue to take place every three years; and reviews of interest rates and eligibility for the Poverty Reduction and Growth Trust will occur every two years. Thematic reviews, on the other hand, take place at Management's initiative or at the Board's request. The Board can ask at any time that any review be undertaken or accelerated.

²⁵Two thematic reviews specified self-evaluation as their primary purpose: “Anti-Money Laundering and Combating the Financing of Terrorism: Review of the Quality and Consistency of Assessment Reports and the Effectiveness of Coordination” (IMF, 2006a) and “Treatment of Exchange Rate Issues in Bilateral Surveillance” (IMF, 2006b).

Table 2. Self-Evaluation in IMF Policy and Other Institution-Wide Reviews: Summary of Findings

Evaluation Questions	Policy Reviews		Other Thematic Reviews		Total	
	Number	Share	Number	Share	Number	Share
Self-evaluation was an explicit purpose of review?	13	57%	6	17%	19	33%
Presented evaluative evidence (e.g., regarding compliance, quality)?	19	83%	19	54%	38	66%
Conducted evaluative analysis to determine effectiveness?	19	83%	18	51%	37	64%
Drew lessons for future policy or operations?	18	78%	12	34%	30	52%
Lessons fed into policy or operational changes?	17	74%	7	20%	24	41%
Included evaluative work pertaining to Staff practices?	14	61%	5	14%	19	33%
Total number of reviews	23		35		58	

instance, the overall theme of the 2011 TSR was “making Fund surveillance as interconnected as the global economy itself,” but it also set out to “take stock of steps taken to address gaps identified in pre-crisis surveillance [and] assess the extent to which surveillance meets stakeholders’ expectations.” Other policy reviews that included self-evaluation explicitly as part of their terms of reference examined, for example, the adequacy of the IMF’s toolkit for low-income countries, the impact of past changes (as well as developments in the global economy) on IMF conditionality and program design, and the effectiveness of the Standards and Codes initiative.²⁶

42. A vast majority of the policy reviews and about half of the thematic reviews (83 percent and 54 percent, respectively) incorporated self-evaluative work as an element of the review, even if they did not have explicit self-evaluative objectives. To varying degrees, these reviews presented, analyzed, and drew lessons from evaluative evidence. About 78 percent of the policy reviews and 34 percent of the other thematic reviews included all of these elements to some extent. Where lessons were identified, these lessons fed into proposed policy or operational reforms in all but one of the policy reviews and a handful of other thematic reviews.²⁷

²⁶“The Fund’s Facilities and Financing Framework for Low-Income Countries” (IMF, 2009a); “2011 Review of the Standards and Codes Initiative” (IMF, 2011a); and “2011 Review of Conditionality—Overview Paper” (IMF, 2012a).

²⁷“Review of the Policy Support Instrument” (IMF, 2009b) concluded that experience was too limited at that time to propose modifications. For many of the thematic reviews for which lessons were not found to immediately having contributed to policy or operational

43. The IEO found many examples of policy and other thematic reviews where the evaluative work was of high technical quality and played a prominent role. Two such examples are the 2011 TSR (IMF, 2011c) and the 2011 Conditionality Review (IMF, 2012a). Both of these reviews analyzed Staff reports and conducted surveys and interviews to examine, respectively, the quality and relevance of IMF advice and the impact of IMF programs on country outcomes. Importantly, the 2011 Conditionality Review included a meta-analysis of lessons from individual programs set out in EPAs and EPEs, although its conclusions about cross-cutting lessons were presented in a background paper and referenced only briefly in the overview report.

44. There was evidence that the integration of self-evaluation in policy reviews helped lessons distilled from experience inform policy development. For instance, evaluative analysis in the 2011 Conditionality Review fueled extensive Board discussion of program design and the application of conditionality; and evidence presented in the 2014 FSAP Review (IMF, 2014d) about limited coverage of nonsystemic countries led to Board discussion about how to enhance attention to financial sector issues in these countries.

45. Further, in many cases, evaluative findings and lessons gained traction in the outcomes of many reviews. For instance, the 2014 TSR finding that about half of Article IV Staff reports did not include an in-depth

reforms, such as the 2009 “Review of Recent Crisis Programs” (IMF, 2009c), these lessons may nonetheless have contributed to future changes.

Box 3. Examples of Self-Evaluation of Staff Practices in Policy Reviews

The 2011 Conditionality Review (IMF, 2012a) and the 2014 Triennial Surveillance Review (IMF, 2014a) are good examples of policy reviews that included substantial self-evaluation of Staff practices, in addition to examining broader institutional policies and operations. Both of these reviews proposed recommendations on how to improve Staff practices, and the TSR also introduced provisions for monitoring implementation.

The 2011 Conditionality Review assessed Staff work in designing programs and engaging with authorities, drawing on evidence collected in desk studies of Article IV Staff reports and surveys of stakeholders. The overview paper highlighted findings about Staff practices such as weak links between program conditionality and the recommendations of prior surveillance. It recommended steps to improve Staff work, such as more consistent discussing of alternative policy options with authorities during program

design. At the Board discussion, Executive Directors reflected on many of these findings and lessons for Staff practices.

The 2014 Triennial Surveillance Review assessed the coverage, depth, and consistency of Staff analysis and the connections between analysis and policy advice in Article IV reports. It identified lessons for Staff practices as well as for the institution's approach to surveillance. For example, it recommended concerted efforts to draw on cross-country policy experiences and to undertake more targeted communications on the analysis of underlying policy advice. In discussing the TSR, Directors addressed a number of these lessons. The Managing Director's Action Plan (IMF, 2014b) set forth mechanisms to monitor implementation by enhancing the existing review process and introducing targeted surveys and informal feedback mechanisms before the next surveillance review.

discussion of the fiscal-monetary policy mix, helped build the case for the decision that these reports should discuss the policy mix more explicitly in order to provide more cohesive policy advice. Similarly, the finding in the thematic review of “Macroeconomic and Operational Challenges in Fragile States” (IMF, 2011b) that the IMF had sought to manage the inherent risks of engaging in fragile states through overly ambitious programs—contributing to uneven program implementation—helped to shape the design of strategy going forward, including emphasis on greater flexibility in program design to better reflect limited implementation capacity. In addition, during a Board discussion in 2012, one Director pointed to the impact of self-evaluation, noting that an increased commitment to honest self-assessment, for instance in the 2009 “Review of Recent Crisis Programs” (IMF, 2009c), was yielding results in terms of improved program design.

46. Policy and thematic reviews also included self-evaluative analysis on Staff practices, although not as often as on institutional issues. About 60 percent of the policy reviews and 15 percent of the other thematic reviews examined aspects of Staff practices, for example, how well Staff handled issues that arose in interactions with authorities on data provision (IMF, 2008 and 2012b). Paying explicit attention to Staff practices is important for improving the execution of IMF policies and activities by Staff teams. **Box 3** provides examples

of self-evaluation of IMF staff practices incorporated in policy reviews.

Strengths and Weaknesses of the IMF Approach to Self-Evaluation of Policy and Thematic Issues

47. The close integration of self-evaluation into the policy development process helped make policy development better informed by lessons from experience, and gave self-evaluation significant traction. However, there were also disadvantages to subsuming self-evaluation in reviews aimed at policy development, particularly that evaluative lessons, especially those focused on Staff practices, could get lost. This was evident in perceptions about reviews and the lack of mechanisms for disseminating their outcomes. First, Staff involved in the preparation of key policy reviews saw policy development as their overriding objective.²⁸

²⁸In the survey for this evaluation, more than half of respondents who had participated in the 2011 TSR agreed that the exercise aimed primarily to develop alternative approaches rather than to look at lessons from past activities, even though nearly all agreed that it also aimed to document what worked and did not work. Only 31 percent of all respondents saw the 2011 TSR as an important accountability tool. The perception that this exercise was more forward-looking than an exercise for learning or accountability was summed up by

Second, Board discussions, and the resulting Summings Up, focused (appropriately) on policy reforms, with little reference to past institutional or Staff performance. Third, there was no established mechanism for disseminating lessons and promoting Staff learning from the evaluative findings in reviews, as the corresponding guidance notes appropriately focused on how to operationalize decisions about policies. Indeed, supporting the concern that evaluative lessons could get lost, nearly half of the respondents for the Staff survey for this evaluation were unaware of whether major reviews of surveillance and conditionality had yielded lessons for Staff and Management.²⁹

one respondent who argued that “success” in such recent reviews seemed to be defined by “the extent to which they [could] be used to promote novelty, in terms of change in the way the Fund works, rather than an empirically based assessment of current practice.” Further, 25 percent of all respondents believed that the 2011 TSR presented predetermined outcomes driven by political factors—making it unlikely that they themselves would rely on the review for lessons from experience.

²⁹Sixty percent of the respondents reported that they did not know whether the 2011 Conditionality Review had contributed to an understanding among Management and Staff about what worked and what did not work; 40 percent of the respondents reported that they did not

48. The recent decision to consolidate some policy reviews and conduct many of them on an as-needed basis, rather than on a preset periodic schedule, has three implications for self-evaluation. First, the consolidation of some policy reviews will widen the scope of these reviews, which will complicate the conduct of in-depth self-evaluation. Second, most periodic reviews (about 60 percent of policy reviews) provided a regular opportunity for assessing the implementation of decisions from earlier reviews. As policy and thematic reviews become less frequent, Staff may need to prepare ad hoc progress reports for this purpose. Finally, for those issues that will be reviewed only on an as-needed basis, there will no longer be periodic opportunities for self-evaluation. This may leave the Board and Management without important information on when changes to policies, operations, or Staff practices may be needed.³⁰

know for the 2011 TSR. Both of these reports had background papers containing important lessons for Staff.

³⁰While it may be justified to undertake policy reviews on an “as-needed” basis when making adjustments in light of changing global conditions or shifts in the views of member country authorities, it is not best practice to assess past performance for learning and accountability purposes on this basis.