

**FOR
AGENDA**

ARCHIVES
ROOM HQ C-532 0414

EBAP/00/29

March 23, 2000

To: Members of the Executive Board

From: The Secretary

Subject: **Review of Experience with Evaluation in the Fund**

The attached paper on the review of experience with evaluation in the Fund, which was approved by the Evaluation Group at its meeting on March 23, 2000, will be brought to the agenda for discussion on a date to be announced.

Mr. Bernes (ext. 37783) and Mr. Chelsky (ext. 34151) are available to answer technical or factual questions relating to this paper prior to the Board discussion.

Att: (1)

Other Distribution:
Department Heads



INTERNATIONAL MONETARY FUND

Review of Experience with Evaluation in the Fund

Prepared by the Evaluation Group of Executive Directors¹

March 13, 2000

Contents

I. INTRODUCTION	2
II. EVALUATION IN THE FUND: THE PRESENT STRUCTURE	3
III. THE PRINCIPLES OF INDEPENDENT EVALUATION	6
IV. FUND EXPERIENCE WITH EXTERNAL EVALUATION, 1996-99	7
V. EXTERNAL VIEWS ON INDEPENDENT EVALUATION IN THE FUND	13
1. <i>Official views</i>	13
2. <i>Public views</i>	14
VI. OPTIONS FOR INDEPENDENT EVALUATION IN THE FUND	16
1. <i>Continue with the Existing Structure</i>	16
2. <i>Expand the Capacity of OIA</i>	17
3. <i>Establish an Independent Evaluation Office (EVO)</i>	17
4. <i>Division of Labour under EVO</i>	19
5. <i>Recommendation of the Evaluation Group of Executive Directors</i>	19
VII. NEXT STEPS	20
VIII. APPENDIX I: TERMS OF REFERENCE OF THE EVALUATION GROUP	22
IX. APPENDIX II : EVALUATION STUDIES 1996-99	23

¹ The Evaluation Group of Executive Directors (EG) is responsible for monitoring the evaluation function in the Fund and advising the Executive Board. It was first constituted in August 1996 with four Directors, and was subsequently expanded to six Directors. At present they are Messrs. Bernes (Chairman), Barro-Chambrier, Eyzaguirre, Mirakhor, Wijnholds, and Yoshimura. The EG is grateful to Mr. Brau (TRE) and Mr. Alexander (OIA) for their valuable assistance and advice in the preparation of this paper.

I. INTRODUCTION

1. This paper reviews the experience with the evaluation of Fund activities in the period 1996-99 with a view to determining how evaluation—and independent evaluation, in particular—can best assist the Fund to carry out its mandate and responsibilities in the future. The paper responds to a request by the Executive Board in June 1996 that such a review should take place. The reasons for the Board's request, and for the authorship of this paper by the Evaluation Group of Executive Directors (EG), are explained in the context of a brief review of the history of the evaluation function in the Fund.²

2. The modalities of Fund evaluation were discussed by the Executive Board on many occasions, including in January 1993 with management's proposal for a separate evaluation office (EVO) that would be independent of the Executive Board, management, and staff. Management's proposal drew on a report prepared in 1992 by a task force of senior staff,³ but there was no consensus in the Executive Board at that time, including on the question of the power of appointment of the Director of an EVO. Many Directors believed that self-evaluation by staff served the needs of the Fund adequately, and many Directors were concerned with the budgetary cost of an EVO and staffing constraints. Establishing an EVO would have required the transfer of experienced staff from operational work at a time when the Fund was under severe pressure from the requirements of many new member countries following the dissolution of the Soviet Union.

3. The issue of independent evaluation was again brought to the attention of the Executive Board in 1995-96 when management, in light of the favorable experience with an evaluation by an outside expert of Fund surveillance in the context of the Mexican crisis, and in an effort to forge a consensus on this issue, proposed an alternative approach. In June 1996 the Executive Board endorsed a "pragmatic approach" to the evaluation function for a trial period.⁴ Rather than establishing an independent EVO, an understanding was reached to continue with existing practices of self-evaluation by the operational departments responsible for the activities being evaluated; to conduct internal evaluations by the Office of Internal Audit and Inspection (OIA) in response to specific needs; and at the same time, to undertake up to 2-3 independent evaluations per year by outside experts under the guidance of the EG. Experience under this approach was to be reviewed after two years. Due to pressures from

² A detailed history of evaluation in the Fund is provided in the background paper authored by the Office of Internal Audit and Inspection (OIA): *Evaluation in the Fund and in Other International Institutions*, to be issued shortly.

³ *Establishing an Evaluation Office in the Fund*, EBAP/92/166 (12/17/92).

⁴ BUFF/96/69, 6/10/96, *Concluding Remarks by the Chairman, Evaluation Function in the Fund – Further Considerations*, EBM 96/55, 6/7/96.

elsewhere in the Work Program and the time taken to arrange for, and conduct, the external evaluations, the review was postponed until the present time.

4. This paper continues, in Section II, with an outline of the existing evaluation structure in the Fund and a list of evaluations undertaken since 1996; Section III discusses the principles which should guide effective independent evaluation; Section IV provides an overview of the external evaluations undertaken by the Fund since 1996 and identifies particular issues arising from the experience that any effort to strengthen the Fund's capacity for systematic evaluation will need to address; Section V examines the views of outside groups—the Interim Committee, the G7 Finance Ministers, and non-government organizations—on how evaluations should be carried out in the Fund; Section VI presents options and a recommendation for Fund's evaluation structure, and Section VII suggests steps that would need to be taken to follow up on the Board discussion.

II. EVALUATION IN THE FUND: THE PRESENT STRUCTURE

5. Since 1996, the Fund has had a tripartite evaluation structure consisting of:

1) Self-evaluation by operational staff: Activities to be evaluated are proposed by management, often after consultation with Directors, and are considered by the Executive Board in the context of the semi-annual Work Program of the Board. Self-evaluations are periodic, and have focussed on a range of issues, including conditionality, surveillance, exchange rate policy, fiscal policy, taxation instruments, monetary policy and financial sector modernization. The results of these evaluations—particularly those related to the core areas of the Fund's work—are presented to Directors and, where appropriate, are discussed by the Executive Board. The publication of self-evaluation reports is determined by management on a case-by-case basis in consultation with the Executive Board.

2) Evaluation by OIA: The topics to be evaluated by OIA are proposed by management, often after consultation with Directors, and may also be considered by the Executive Board in the context of the Work Program. In accordance with the organizational rules of the Fund, OIA evaluation reports are provided to management; they have then been issued without change to the Executive Board for information or discussion. As with self-evaluation reports, the publication of OIA evaluation reports is determined by management on a case-by-case basis. In May 1996, the staff allocation of OIA was expanded by two economist positions to enable it to undertake a small number of internal evaluations and to assist the EG in the oversight of external evaluations. Staff who had not been involved in the activities being evaluated were seconded to OIA for a 1-2 year period. Subsequently, research assistant and consultant resources were added to OIA. Three evaluations have been carried out by OIA since that time—of the resident representatives program, technical assistance activities, and the provision of general services in the Fund. These reports were not published.

3) Independent evaluation by outside experts: In accordance with the understanding reached by the Board in June 1996, the EG makes a recommendation to the Executive Board on topics for external evaluation, including their scope, terms of reference, the choice of outside evaluators, and any other relevant aspects of the evaluation.⁵ Following the initiation of an external evaluation, the EG has issued a public statement providing a general description of the evaluation undertaken and its intended timetable. The publication of external evaluation reports is determined by the Executive Board on a case-by-case basis. Three such evaluations—of the ESAF, surveillance, and research activity in the Fund—have been completed. All were published shortly after Executive Board consideration.

6. The first category – self evaluation – is an endogenous exercise and, in any large organization like the Fund, is integral to ensuring the institution’s ongoing effectiveness and efficiency. By its very nature, it entails a high degree of ownership of the conclusions and recommendations reached. Fund experience with self-evaluation has a long history and has contributed importantly to making the Fund a more effective organization. The results of self-evaluations have frequently been published and have thereby contributed to increased public understanding of the Fund’s work. The continuation of such efforts is non-controversial; self-evaluation should remain an integral element of the Fund’s operations. The more substantive issue is with respect to how the benefits of self-evaluation can be enhanced by various forms of *independent* evaluation. The remainder of the paper will focus on this question, including the issue of how agreed-to independent evaluation efforts should be coordinated with self-evaluation to determine how to deliver the maximum benefit to the Fund and its members in the most efficient manner.

⁵ *The Terms of Reference of the Evaluation Group, which were approved by the Executive Board in September 1996 are shown in Appendix I.*

Box 1. Evaluation Studies Conducted During 1996-99

- Self-evaluations:**
- (i) 1997 Biennial Review of Surveillance ¹ (PDR)
 - (ii) Review of the ESAF ² (PDR)
 - (iii) Lessons for Surveillance from the Asian Crisis ³ (PDR and FAD)
 - (iv) Review of Fund-Supported Programs in the Asian Crisis ⁴ (PDR)
- OIA Evaluations:**
- (i) Resident Representatives Program ⁵
 - (ii) Technical Assistance Activities ⁶
 - (iii) General Services Review ⁷
- External Evaluations:**
- (i) MAE technical assistance ⁸
 - (ii) Program design in European transition countries ^{9c}
 - (iii) ESAF ¹⁰
 - (iv) Research ¹¹
 - (v) Surveillance ¹²
 - (vi-ix) External communications ¹³

¹ *Biennial Review of the Implementation of the Fund's Surveillance: SM/97/53 (2/19/97), SM/97/92 (4/10/97), and EBM/97/24 (3/14/97).*

² *Review of Experience Under ESAF-Supported Arrangements: EBS/97/112 (6/23/97), EBS/97/123 (7/2/97), and EBM/97/75 (7/21/97).*

³ *Review of Members' Policies in the Context of Surveillance - Lessons for Surveillance from the Asian Crisis: EBS/98/44 (3/9/98), and EBM/98/34 (3/26/98).*

⁴ *Fund-Supported Programs in the Asian Crisis: EBS/98/202 (11/25/98), and EBM/98/130 (12/18/98), and Summing-Up by the Chairman (revised): BUFF/98/117.*

⁵ *Review of the Resident Representative Program: EBS/97/137 (7/25/97), and EBM/98/7 (3/2/98).*

⁶ *Review of Fund Technical Assistance: EBAP/99/59 (5/17/99), and EBM/99/61 (6/8/99).*

⁷ *General Services Review: EBD/99/52 (4/9/99).*

⁸ *External Evaluation of Technical Assistance Provided by the Monetary and Exchange Affairs Department Report of Independent Panel: EBS/96/15 (1/26/96), and EBM/96/47 (5/15/96).*

⁹ *WP/96/108, WP/96/125, and WP/97/31. A Board paper was not issued for this evaluation.*

¹⁰ *External Evaluation of the Enhanced Structural Adjustment Facility: EBAP/98/8 (1/22/98), and EBM/98/25 (3/11/98).*

¹¹ *External Evaluation of the Fund's Economic Research Activities: EBAP/99/85 (7/15/99), and EBM/99/99, (9/7/99).*

¹² *External Evaluation of Fund Surveillance; EBAP/99/86 (7/15/99), and EBM/99/94 (8/27/99).*

¹³ *Four studies sponsored by EXR (see Annex) are summarized in Strengthening the Fund's External Communications—Plans and Resource Implications; SM/00/14 (1/27/00).*

III. THE PRINCIPLES OF INDEPENDENT EVALUATION

7. In order to establish a framework to assess the potential contribution of *independent* evaluation to the Fund's work, as well as the appropriate modalities for such evaluation, it is helpful to spell out why an institution staffed, as is the Fund, with such highly-skilled individuals would seek to augment its capacity for self-evaluation with more independent forms of evaluation. With clarity on the underlying motivation, it then becomes easier to articulate guidelines for effective independent evaluation against which the Fund's existing capacities and practices can be measured.

8. As with self-evaluation, the value of independent evaluation at the Fund—and indeed, any large organization—is the opportunity it provides to reflect on past efforts in an attempt to identify patterns and interlinkages which can be used for one or all of the following purposes:

- to provide additional objective information to assess the performance of an organization's activities (the results assessment purpose);
- to improve an organization's activities through feedback of lessons learned (the learning purpose); and
- to provide accountability to the organization's shareholders and the public for the results of its activities in the absence of market criteria by which to measure its effectiveness (the accountability purpose).⁶

9. From an administrative and organizational perspective, independent evaluation provides the opportunity for review where staff directly involved in the activities in question lack the time to systematically assess their own efforts, or may not be positioned to adequately situate their work in the context of other broader institutional efforts and practices. Even the *perception* that evaluation is independent adds value, including by enhancing the external credibility of the evaluation's conclusions and, by extension, of the institution itself. As such, independent evaluation can serve as a useful vehicle to inform the work of critics, academics, and policy analysts outside the institution. In the Fund's case, this has the potential to assist in generating more broadly-based public support for its work.

10. The motivation for particular independent evaluations differs from evaluation to evaluation. This could include the desire to enhance operational efficiency and/or policy effectiveness, the bringing to bear of expertise that may not be available within the institution, or the enhancement of external confidence in the work of the institution. Given the specific motivation for a particular evaluation, its efficacy will vary depending on a variety of factors, including:

⁶ This purpose is distinct from accountability for use of public funds in an accounting and legal sense, which is the responsibility of an auditing office.

- the evaluators' level and breadth of experience with the subject matter;
- their familiarity with the institution's internal processes, corporate culture, and history;
- their actual and perceived impartiality;
- their independence from the process of policy involving line management and operations;
- their access to both formal and informal sources of information;
- the quality of their reputation with those whose work they are evaluating, with managers, and with relevant third parties outside the institution;
- the rigor of the methodology they apply;
- the extent to which management/decision makers take ownership of the recommendations arising from the evaluation; and
- the existence of a mechanism to ensure that the conclusions of the independent evaluation are appropriately integrated into the work of the institution.

11. Consideration of these factors raises questions about how independent evaluation should be undertaken and who should be tasked with the work; the relationship of the evaluators to those being evaluated and to the institution itself has a significant impact on the evaluation's outcome as well as its effectiveness in meeting chosen objectives. Should, for example, evaluators be employees of the institution? If so, what should be their relationship to the institution's managerial and governing structure so as to maintain their independence? What role can and should be played by individuals and groups outside the institution?

12. While it is possible to articulate broad principles with which to govern the choice of evaluators and evaluation modality, the most appropriate modality—one that takes into account both independence and efficiency concerns—will vary depending on the subject being evaluated and the motivation for the evaluation itself. This suggests that no particular modality can *a priori* be considered superior to other modalities; each will have its own strengths and weaknesses. Therefore, any structure intended to enhance the Fund's capacity to undertake effective independent evaluation will need to be flexible enough to adapt to the demands of, and motivations for, particular evaluations.

IV. FUND EXPERIENCE WITH EXTERNAL EVALUATION, 1996-99

13. The pilot project on external evaluation launched in 1996 represented one possible modality. This section reviews various aspects of that experience and, in light of earlier discussions with Directors, staff, management, and the evaluators themselves, discusses their strengths and limitations. Section V draws on this discussion to present options for enhancing the Fund's capacity to undertake systematic independent evaluation. All told, there were three external evaluations motivated by the EG and undertaken in the period from

1996 to 1999. These covered a range of topics and were of different scale—1) ESAF (1996-97); 2) Surveillance (1998-99); and 3) Research (1998-99)⁷

14. Choosing Topics: The decision to undertake an external evaluation of ESAF was reached simultaneously with agreement on the pilot project. There was broad and early agreement that the ESAF was an appropriate first case for the pilot for a number of reasons, including the timing of the internal evaluation of ESAF, the desire to extend the coverage of the internal evaluation into areas for which the Fund had limited experience (e.g., the social sector implications of ESAF programs), and the need to build support for ESAF in member countries to facilitate the generation of resources to finance a self-sustained ESAF.

15. The process of achieving consensus on subsequent external evaluation topics was—as might be expected given the learning process inherent in the approach—somewhat more time consuming. The EG met in June 1997 to discuss topics for possible external evaluations. Many of the topics had been suggested in the context of the Board's regular discussions, including of the Work Program. A variety of topics were suggested, including surveillance, exit strategies from managed exchange rate regimes, financial sector reforms, status of the reform process in transition economies, institutional aspects of the reform process, the use of statistics, and research activities. Directors proposing specific topics were asked to prepare short notes describing the reason for, and scope of, the topics they were advocating. In October 1997, one major evaluation (Surveillance) and one smaller evaluation (Research) were agreed upon.

16. Subsequent to the agreement to undertake these two evaluations, the EG met again to discuss additional evaluations. Prior to this meeting, and in the context of the discussion of the Administrative and Capital Budgets for FY2000 (EBM/99/45, 4/20/99), a group of ten Executive Directors requested the initiation of an external evaluation of the Fund's internal processes and procedures. This topic was discussed, along with others, at the October 1999 meeting of the EG and received broad endorsement by Directors in the context of the January 2000 discussion of the medium-term budget outlook. Modalities for such a review are being considered. Around the same time, EXR initiated four external assessments of various aspects of the Fund's external communications strategy and effectiveness (see Annex footnote No. 27), the results of which were reported informally to Executive Directors.

17. Under this arrangement, the experience in choosing topics was a broad-based one and one that ensured that the topics chosen were of interest to a large cross-section of the Board. On the other hand, achieving consensus on individual topics was time and resource intensive, and considerable time spent by individual Directors preparing proposals for individual topics

⁷ In addition, two other external evaluations were undertaken in 1996 -- on aspects of program design in European transition countries and on MAE technical assistance. These were sponsored by EUI and EUII and MAE, respectively, prior to the establishment of the EG.

and by staff in responding to, and analyzing, the various requests advanced over a period of five months. Also, many topics which had significant minority interest but were either not priorities for the majority of Directors or were not of interest to the full membership were not undertaken.

18. Terms of Reference: In all three cases, proposed terms of reference were drafted by staff on the basis of prior discussion within the EG. These draft terms of reference were subsequently discussed and developed by the EG. This too was a time consuming process as Directors sought to articulate terms of reference that were clear and coherent, while at the same time, did not excessively prescribe the content of the evaluation so as to overly constrain the evaluators. An interesting observation made by one member of the EG, in the context of the ESAF evaluation, was that excessive effort to precisely articulate the terms of reference was not a productive exercise since evaluators of the caliber chosen would themselves decide what questions were relevant to the overall topic. For the chosen evaluators to ignore some aspects of an issue that they considered important because Directors did not want those issues addressed would be seen as undermining the independence of the evaluation and the reputation of the evaluators.

19. Choosing the External Evaluators: Directors articulated broad principles to guide the choice of evaluators for each evaluation (e.g., balance in regional representation, professional characteristics and experience). After consulting with management and other Directors, the Chairperson of the EG put forward recommendations to the EG.

20. The selection process was generally uncontroversial, although the desire to ensure broad and balanced geographic representation on each evaluation team complicated the selection process somewhat. In the case of the research evaluation, the EG explicitly sought evaluators with strong academic credentials. It is noteworthy that, in the context of the subsequent discussion of this report, a few Directors and staff were critical of a perceived "overly-academic" approach taken by the evaluators. Some staff criticized the evaluators, particularly in the context of the evaluation of surveillance, for a perceived lack of familiarity with the Fund's mandate and operations. This points to the challenge inherent in trying to balance the importance of "fresh" perspectives on Fund issues while at the same time ensuring adequate familiarities with the Fund and its operations.

21. Staff Involvement in Pilot Project: Consultations with staff naturally figured prominently in the preparation of all external evaluations. Staff in area and functional departments spent time assembling information for the evaluators. OIA staff resources were also drawn on to provide administrative and research/technical support to the external evaluators (approximately 0.9 staff years for the ESAF evaluation, 0.8 staff years for the surveillance evaluation, and 0.4 staff years for the research evaluation).

22. In all cases, staff were given the opportunity to provide comments—both factual and substantive—on a draft of the evaluation report. With respect to the final version of the

reports, staff prepared formal responses for consideration by Executive Directors. These responses were published along with the evaluation reports themselves.

23. No formal guidelines were established at the start of the pilot project on when staff would be provided with the final evaluation report or when the staff response would be made available to Directors. In the case of the ESAF evaluation, the final report was distributed to staff, management and Directors at the same time; the staff response was issued some 2-3 weeks later and a little over one week before the Board discussion. For both the surveillance and research evaluations, staff and management were provided with the final evaluation reports before they were seen by Executive Directors to permit the simultaneous circulation of staff responses and the final reports to Directors.

24. An innovation introduced for the research and surveillance evaluations was to hold an informal session prior to the formal discussion of each evaluation to provide Directors with an opportunity to ask questions of, and seek clarification from, the evaluators on the content of the evaluation report prior to expressing formal views on the recommendations themselves. The Chairman of the EG had considered having staff delay distribution of their response until after the informal session to ensure that the focus of the session was the content of the evaluation report itself rather than both the evaluation and the staff response. It was also thought that this would help ensure that the informal session focused on clarifying the content of the evaluation prior to a discussion of the merits of its conclusions. This delay in the circulation of the staff response was strongly resisted by some staff and management who argued that the staff response should be provided to Directors *at the same time* as the evaluation reports lest Directors begin to form views on the reports' content without the benefit of staff input. In the end and as noted, staff were permitted to distribute their comments simultaneously with the staff report and *prior to* the informal information session with evaluators. No formal policy has yet been established leaving the decision on the appropriate sequencing to be determined on a case-by-case basis.

25. Management's Involvement: With the ESAF evaluation, there was no formal management response to the evaluation. In the cases of the surveillance and research evaluations, and at the request of some Executive Directors, management prepared its own formal response to the evaluation reports. These statements were distributed a few days prior to the formal Board meeting and well after the informal session with evaluators. Both management statements were published along with the final report on the Fund's website. A statement from management separate from that of staff was welcomed by many Directors during the formal Board discussions, particularly as the statements focused on broader issues that transcended the specific topics.

26. Follow-Up to the Evaluations: One of the challenges evident in the pilot project was ensuring systematic follow-up of evaluations in the midst of intense pressure on Executive Directors' time and high turnover in the Board. Despite the fact that a paper entitled "Distilling the Recommendations of the ESAF Evaluations" was prepared subsequent to the

Board discussion of the ESAF Evaluation (which presented an action plan to incorporate the recommendations of both the external and internal (PDR) evaluations of ESAF), a number of Directors and individuals outside the institution were concerned that there did not appear to be systematic monitoring of the implementation of this action plan. This led a number of Directors to request staff to prepare a short note to take stock of efforts underway to respond to the recommendations of the ESAF Evaluation; "Status Report on Follow-Up to the Reviews of the Enhanced Structural Adjustment Facility" (EBS/99/173, 8/30/99) was prepared and posted on the Fund's website.

27. To avoid the emergence of public skepticism with respect to the follow-up to independent evaluation, an effort was made to formalize the follow-up process for both the research and surveillance evaluations. Directors requested that staff and management prepare an action plan responding to the recommendations of each evaluation and that this be followed by a formal stock-taking effort approximately one year after the discussion of the reports by the Board. The Summing Up of the discussion of the surveillance evaluation indicated that "Management intends to come back to the Board after the Annual Meetings with precise suggestions on a program to deal with the issues raised by the External Evaluation Report. These issues will also be followed up in more detail in the Biennial Review of Surveillance scheduled for end-1999"⁸. To date, no precise program has been brought forward and the Biennial Review of Surveillance has been delayed from end-1999 and is now scheduled to take place in March 2000.

28. No action plan was presented by management to Executive Directors to follow-up on the external evaluation of research. However, a number of the recommendations endorsed by the Board have been implemented, namely, the establishment of the Committee on Research Priorities and drawing up a RES action plan. Included in the latter are more frequent participation by RES staff in external conferences, an annual research conference at the Fund, a joint research seminar series with the World Bank, and a new research newsletter to disseminate Fund research to non-technical audiences.

29. External Relations and Communications: The Fund's approach to external communications evolved throughout the pilot project. At the time the terms of reference were agreed for the ESAF evaluation, Directors engaged in an in-depth discussion of the appropriate communications strategy. Some Directors preferred no public announcement be made concerning the initiation of the Evaluation. They were concerned that the evaluators would be targeted by special interest lobby groups should their endeavors become widely known. It was their view that institutional silence on the evaluation would permit the evaluators to work in "peace and serenity".

⁸ "Summing Up by the Chairman: Report of the External Evaluators on Fund Surveillance", Executive Board Meeting 99/100, SUR/99/109, September 10, 1999

30. Other Directors advocated full transparency concerning the initiation of the evaluation, its terms of reference and the names of the evaluators. They considered the initiation of an External Evaluation of ESAF to be a positive development which could help counteract the impression that the Fund was secretive and not open to criticism from outside. In the end, a compromise was reached in which the initiation of the ESAF Evaluation was announced in a press release but the names of the evaluator's were omitted. After a period of months, the evaluators returned to the Fund and requested that their names be formally released to the public, arguing that secrecy was unrealistic given the consultations being undertaken in the context of the evaluation. Since this time, the public statement issued at the initiation of each external evaluation has contained the names of the evaluators chosen.

31. In the conduct of their work, evaluators were given complete freedom of choice in determining with whom they would consult. No constraints were placed on their ability to conduct their studies in an open and public manner nor were evaluators explicitly required to solicit input from particular non-official groups, including civil society, broadly defined. Suggestions as to whom the evaluators may want to consult were, however, contained in the terms of reference for each evaluation but the ultimate decision on the scope of consultations was left to the evaluators' discretion.⁹

32. With respect to the results of the evaluations, it was agreed that no *a priori* decision would be made to publish the reports or any associated documentation. Instead, this decision would be made by Directors on a case-by-case basis after reviewing the results of each evaluation. In practice, however, all reports were published on the IMF's web site unedited (with the exception of factual corrections), along with associated documents including staff comments, the Chairman's Summing Up and a statement by the Chairman of the EG. Summaries of all three evaluations were contained in IMF Annual Reports¹⁰. Moreover, for

⁹ *The Terms of Reference for the ESAF Evaluation indicated that "at their full discretion, the evaluators may wish to take into account the views of concerned country authorities and social partners; of parliamentarians; of representatives of multilateral development banks, bilateral donors and non-governmental organizations; of academic experts; and of Fund Executive Directors and staff". For the Surveillance evaluation, it was stated that "at their full discretion, the evaluators may wish to take into account the views of member country authorities, parliamentarians, academic experts, representatives of other international organizations, representatives of the business and financial market communities, representatives of civil society and the media, and Fund Executive Directors and staff". For the Research Evaluation, the illustrative list included, member country authorities, academic experts, representatives of other international organizations, and Fund Executive Directors and staff.*

¹⁰ *The publication record of other independent evaluations was mixed. With regard to the evaluations by OIA, none of the reports were published, but a short summary of the review of resident representatives was included in the 1998 Annual Report, and of the general services review in the 1999 Annual Report. The evaluations on aspects of program design in European transition countries was published in the form of a series of Working Papers. There was no publication of the external evaluation of MAE technical assistance although a summary of the report was contained in the 1997 Annual Report.*

each of the evaluations, open public fora and/or press conferences were hosted by the Fund, with the participation of staff and management, at which the evaluators presented the results of their work and responded to questions from individuals and the media.

V. EXTERNAL VIEWS ON INDEPENDENT EVALUATION IN THE FUND

1. Official views

33. Recent communiqués of the **Interim Committee/International Monetary and Financial Committee** have noted the importance of evaluation in contributing to the transparency of the Fund. The Communiqué of October 4, 1998 noted that greater openness about the Fund's own policies and the advice it provides to members should be "strengthened through ... more public information on, and evaluations of, the Fund's operations and policies." This view was reaffirmed in the Communiqué of April 27, 1999. The Communiqué of September 26, 1999, clarified the type of evaluation to which the Interim Committee was referring:

"The Committee welcomes the recent independent, external evaluations of IMF surveillance and research activities and encourages the Executive Board to examine the recommendations of the former further in the context of the next internal review late in 1999. The Committee also reaffirms the importance of independent evaluations of the Fund's operations and policies" (underlining added).

34. The importance of independent evaluation to strengthening the international financial architecture has for some time been noted by the **G7 Ministers of Finance**. Their first recommendation—to establish an independent evaluation office at the Fund—was made in the documentation for the June 1995 Halifax Summit.¹¹ For the May 1998 Birmingham Summit, the G-7 Finance Ministers called on the Fund to explore ways to make external evaluation "more systematic".¹² Similarly, in October 1998, G7 Finance Ministers and Central Bank Governors called for:

"The IMF to develop a formal mechanism for systematic evaluation, involving external input, of the effectiveness of its operations, programs, policies, and procedures"¹³

¹¹ *The Halifax Summit Review of the International Financial Institutions – Background Document, June 1995.*

¹² *Strengthening the Architecture of the Global Financial System – Report of G7 Finance Ministers to G7 Heads of State or Government for their meeting in Birmingham, May 1998.*

¹³ *Declaration of G7 Finance Ministers and Central Bank Governors, October 30, 1998.*

35. In June 1999,¹⁴ G-7 Finance Ministers agreed to take steps to improve the effectiveness of the IMF and other international financial institutions by, *inter alia*, “encouraging the IMF to continue undertaking systematic evaluations both internal and external of the effectiveness of selected operations programs, policies, and procedures”.

2. Public views

36. The Center of Concern, a Washington DC-based NGO focusing on international social issues, organized an “IMF Study Group” in June 1997 to consider the issues of transparency and evaluation at the Fund¹⁵. The Group was convened at a time when Fund policies were being questioned by a broad spectrum of public opinion. The Group worked with the Chairman of the EG, senior IMF staff from operational and the External Relations departments, representatives of developing country governments, experts from universities and multilateral development banks, representatives from non-governmental organizations critical of the Fund, and staff members of U.S. Congressional offices.

37. In July 1998, the Chairman of the EG invited the Center for Concern to discuss its report¹⁶ with the EG and Fund staff at a meeting at headquarters. Their report noted that systematic evaluation of Fund activities was indispensable in promoting transparency as it would provide information on the Fund’s objectives, the terms of its lending programs, and the outcome of these programs. It outlined the principles that should guide evaluation— independence of the evaluators, effectiveness and transparency of the evaluation process, and comprehensiveness of the scope of evaluation. It recommended the continuation of existing self-evaluation mechanisms, and the addition of an independent evaluation unit structured in accordance with these principles. It noted that “a separate evaluation office that is carefully structured to be, to the fullest extent possible, independent from management and the Executive Board can establish a reputation with the outside world that its reports are indeed, objective, and can thus contribute more to the confidence of the public at large in the institution than can be achieved by any internal units, however capable and independent-minded its staff.”

38. The report’s more detailed recommendations were:

- The independent evaluation office should be staffed primarily from Fund staff members whose careers should be assured.

¹⁴ *Strengthening the International Financial Architecture – Report of G7 Finance Ministers to the Köln Economic Summit – Cologne, 18-20 June, 1999.*

¹⁵ *The Group comprised 14 members including noted economists Albert Fishorn and Peter B. Kenen, and former senior IMF staff members Aziz Ali Mohammed and Jaques J. Polak who authored the Group’s report.*

¹⁶ *Jaques J. Polak, IMF Study Group Report: Transparency and Evaluation; Report and Recommendations by a Special Study Group convened by the Center of Concern, April 1998.*

- The independent evaluation office should be able to use outside experts to evaluate aspects of programs, such as their social effects, for which it does not have the expertise itself and could not economically hire the expertise on a permanent basis.
- The Fund should not find it necessary to also have recourse to panels of outside evaluators appointed by the Executive Board. The report suggested that there might be a conflict of interest in that the Executive Board, which was ultimately responsible for the policies that the evaluators are charged to appraise, would not approve topics for evaluation that would reflect unfavorably on the role of the Board itself, or that experts known for their “unfair criticism” of the Fund would not be chosen as evaluators.

39. Another report was published on the issue of independent evaluation at the Fund in April 1998 by two NGOs—Friends of the Earth and the Rethinking Bretton Woods Project. Representatives of these Groups were also members of the Center of Concern’s Study Group and attended the July 1998 meeting with the EG. While endorsing many of the recommendations of the Study Group, the views of these representatives diverged on some issues.¹⁷ Chief among the differences was the need they perceived for greater public participation in the evaluation process. This followed from their underlying premise that, since Fund programs and advice directly affect all strata of society in member countries, the Fund should be accountable to civil society as a whole, and not only to governments. They also expressed concern with the absence of staff dedicated to the evaluation process which had resulted in infrequent reviews and the absence of adequate follow-up to the recommendations of particular evaluations. While they acknowledged that the Directors provided oversight to the conduct and follow-up of evaluations, they considered Directors to be too busy to shepherd the evaluation process on a timely basis.

40. In light of the perceived weaknesses in the existing evaluation structure in the Fund,¹⁸ the report concluded that independent evaluation of Fund operations was needed to better direct financial resources to effective programs. Four options were proposed for consideration:

¹⁷ Angela Wood and Carol Welch, *Policing the Policemen – the Case for an Independent Evaluation Mechanism for the IMF, Rethinking Bretton Woods Project and Friends of the Earth US, April 1998.*

¹⁸ *These weaknesses were identified as: (i) lack of objectivity of self-evaluations; (ii) the scope of internal reviews was limited by the framework in which the Fund operates – “issues of importance to civil society such as social impacts of programs and the degree of local ownership of a program are unlikely to be chosen as (evaluation) topics;” (iii) lack of information on the extent to which Fund management follows through on the findings and recommendations of internal reviews; (iv) lack of transparency in some external evaluations – there was no guarantee that findings will be published; (v) lack of expertise in self-evaluations to examine the impact of Fund programs on poverty, different socioeconomic groups, and the environment; and (vi) lack of assurance that the existing interdepartmental review procedure and the review function of the Executive Board in the Article IV consultation process will result in lessons learned being included in new programs.*

- maintaining the present evaluation structure but allowing public participation in decisions regarding the choice of topics to be evaluated, the selection of outside experts, and framing the terms of reference. They also argued that staff with a wider range of skills would need to be hired to improve the quality and scope of self-evaluations;
- establishing an independent in-house evaluation unit, similar to that in the World Bank, of sufficient size to be able to evaluate on-going programs. The initial cost of such a unit was acknowledged as a constraint, but one which should be balanced against the longer term savings for the Fund arising from more effective programs and the additional resources likely to result from greater public support;
- establishing a common evaluation unit for both the Fund and the World Bank which would be staffed independently. However, because of this separation, there would be some difficulty in ensuring the absorption of recommendations into operations; and
- establishing an evaluation committee for the Fund comprised of external experts and possibly a retired senior IMF staff member to help in disseminating recommendations within the Fund. This committee would oversee external evaluations but would not be able to carry out continuing evaluation of completed programs.

VI. OPTIONS FOR INDEPENDENT EVALUATION IN THE FUND

41. Drawing on the experience over the past three years with self-evaluation, internal independent evaluation and external evaluation, including in the context of the Pilot Project, as well as external views on the conduct of evaluation in the Fund, this section presents three options for the future, discusses their advantages and disadvantages, and advances a single recommendation for proceeding.

1. Continue with the Existing Structure

42. *Self-evaluations by operational departments, a limited number of internal evaluations by OIA staff, and occasional external evaluations by the EG:* In defense of the status quo, it should be noted that the higher profile of evaluation in the Fund has encouraged the institution to be more self-critical; and the publication of the three most recent external evaluation reports and of several self-evaluations, together with follow-up action in the case of the internal and external ESAF evaluations,¹⁹ has improved the public's understanding and acceptance of the Fund's operations. At the same time, it has also generated the expectation from both the official sector and civil society that the Fund will continue to be subjected to regular independent evaluation.

43. However, the current approach has a number of shortcomings. Self-evaluations continue to suffer from the perceived partiality of those undertaking the analyses. This might

¹⁹ A Status Report on Follow-Up to the Reviews of the ESAF prepared by PDR and FAD was published on the Fund's external web site on August 30, 1999.

also be perceived to be a shortcoming of expanding the conduct of independent evaluation through OIA given the fund's governance structure and the Director of OIA's ultimate accountability to management.

44. With respect to the external evaluations, there has been criticism of the small number of evaluations undertaken, the quality of some of the analyses, and the practical value of some of the recommendations. A perception might also develop that the direct involvement of the Board in the choice of topics and evaluators could constrain the extent to which sensitive topics are chosen and critical perspectives are brought to bear. Also, the *ad hoc* nature of the external evaluation process presents challenges for the adequacy and transparency of follow-up to the recommendations contained in the evaluations and makes it more difficult to maintain institutional memory of the experience obtained in particular evaluations. There have also been calls from outside groups to play a greater role in various aspects of the external evaluation process, including through consultation on the substance of reviews and in the choice of topics for review. The resource requirements of the present structure, while hard to measure accurately in staff, Board and evaluator terms have been substantial.

2. Expand the Capacity of OIA

45. *Build on the status quo by expanding OIA to make it possible to evaluate selected surveillance work and Fund programs on a continuous basis, to conduct general evaluations of some Fund activities, and to systematize the follow-up to specific evaluations.* Beyond the issues articulated in Option 1, this option would build a foundation of case studies on surveillance and use of Fund resources to augment the information derived from self-evaluations and to constitute an information base for external evaluations. The ongoing existence of OIA would facilitate systematic follow up to the recommendations of particular evaluations. However, as noted above, a shortcoming to this approach remains that the evaluations would not be, or be perceived to be, truly independent given the OIA's accountability to management. This would be further compounded by the competing responsibilities in areas where OIA is most active.

3. Establish an Independent Evaluation Office (EVO)

46. *Establish an EVO at the Fund, reporting directly to, but operating at "arms length" from, the Board, and with effective independence from management.* The EVO would be headed by a Director selected by the Executive Board, in consultation with the Managing Director. Upon appointment, the Director would have full independence from management and staff. The Executive Board (through the EG) would be responsible for ensuring the effectiveness of EVO, implying that the Director of EVO would therefore need to be

ultimately accountable to the Board²⁰. However, the need to ensure the effective and perceived independence of EVO would require it to operate with effective independence from the Board requiring careful drafting of the accountabilities and expectations of the EVO and the EG. At the same time, Directors would need to be adequately informed on the work of EVO to ensure the efficacy of independent evaluation in the Fund.

47. The Executive Board, on advice from management and a newly-mandated EG, would set the budget for EVO and could—throughout the year and in the context of discussions of the Work Program—recommend topics to be addressed. Neither the EG nor the Board, however, would be able to prevent topics from being evaluated or influence the content of evaluations conducted by the EVO. Executive Directors would still retain the right to launch external evaluations themselves, with logistical support from the EVO; this would most likely occur in the context of discussions of the EVO's work program and/or other appearances of the Directors of EVO before the EG.

48. Under this structure, OIA would continue to exist but with a more focussed mandate. Fund departments would continue to conduct periodic self-evaluations as required.

49. As with self-evaluation and OIA evaluation, EVO evaluations would be of value not only in themselves in identifying lessons for the future, but also as an information resource for outside experts conducting any external evaluations undertaken by the Board. At the same time, the structure of EVO would permit it to systematically and consistently monitor the implementation by management and staff of recommendations from prior evaluations, and ensure that follow-up was not diverted by other demands on Directors' attention or frequent turnover in the Executive Board.

50. The effectiveness of this approach would depend crucially on the skill and independence of the Director of EVO, the extent to which qualified and motivated individuals would see work with EVO as rewarding both in and of itself and with respect to longer-term career prospects. Great care would need to be taken to ensure that the EVO maintains its character as an informed and independent evaluators of Fund work and that it not come to be seen as a simple extension of Fund operations. Given the likely costs associated with the creation of EVO, its existence must also be warranted not just on its "public relations" merits but also on the basis of the substantive contribution it makes to enhancing the effectiveness and efficiency of the Fund.

²⁰ *Membership in the EG would need to be left in the hands of the Executive Directors rather than management.*

4. Division of Labour between OIA, EVO, and Self-Evaluation

51. If an EVO is created, there are a number of administrative and technical issues which will need to be addressed. Among these is the need to articulate the appropriate division of labour between EVO, OIA and the self-evaluation by operational departments. As noted, the value of independent evaluation is measured by the extent to which it complements existing evaluation efforts. Therefore, key to the success of the establishment of EVO will be the extent to which Fund departments, including OIA, and EVO coordinate their efforts. While some degree of overlap in topics and functions reviewed is both appropriate and desirable, given the different perspectives which each unit can bring to bear, it will be important to ensure that the institution as a whole allocates its resources in a manner roughly in line with their relative importance. In this regard, managerial and Executive Board guidance will be helpful, as will regular contact between operational departments, OIA, and EVO.

52. Operational departments would continue to undertake reviews of policies (surveillance, conditionality), use of Fund resources programs, country surveillance work, technical assistance programs, and financial policies and activities. OIA—which would continue to report to management—would have primary responsibility for reviews of administrative activities and organizational issues (aimed at enhancing the effectiveness and efficiency of the functioning of day-to-day operations Fund-wide), and share the responsibility for audits of financial statements with the Fund's external auditors. In addition, OIA would take the lead in implementing a comprehensive assessment of the risks faced by the Fund in the conduct of its economic, financial, and administrative activities.

53. In principle, EVO would be unconstrained in its choice of evaluation topics but would be guided by the desire to provide value-added to the Board's work. Evaluations could take the form of broad cross-sectional reviews of policy effectiveness and implementation or of in-depth studies of specific programs in particular countries. In carrying out its mandate, EVO would have complete freedom to solicit external input in the conduct of its activities. It would also be expected, with the assistance of EXR, to undertake whatever external communications activities it would deem necessary to establish and enhance its credibility.

54. The Executive Board would still retain the ability to initiate purely external evaluations (i.e., choose the topic and evaluators and the set terms of reference) but the existence of the EVO would make the need for such effort rare. Were the Executive Board to decide to undertake such an effort, it would—to the extent appropriate—receive administrative support from EVO.

5. Recommendation of the Evaluation Group

55. Self-evaluation at the Fund is widely-perceived to be of high quality. Any extension of the Fund's evaluation capacity must clearly be of the same high quality. At the same time,

it must complement existing evaluation efforts by augmenting the potential scope of evaluation where Fund expertise may be limited and it must enhance the credibility of evaluations to observers outside the Fund. Further, any measures taken to enhance the Fund's evaluation capacity will need to include a transparent and efficient mechanism for systematic follow-up. In this regard, Option 3 would seem the most effective way to achieve the broad range of objectives.

56. Where Option 3—with its emphasis on independence from management and the Board—can improve the existing structure most strongly is in strengthening the credibility of Fund analysis with constituencies outside the Fund (both official and non-governmental). Even if it were internally accepted that current self-evaluation was wholly objective, the perception outside the institution that such bias exists, in and of itself, undermines the ability of the Fund to undertake its work.

57. Given that EVO may legitimately undertake evaluations from a range of perspectives, work of the EVO would benefit from the hiring of staff embodying considerable breadth in their backgrounds and expertise. For example, while all staff would need to have adequate understanding of the macroeconomic issues core to the Fund's mandate, EVO staff should collectively possess both a broad and demonstrated interest and the experience in areas such as public policy, law, economic history and capital markets. The build-up and retention of evaluation expertise in the EVO would also benefit the Fund.

58. For EVO to be effective, it would need to be large enough to carry out and follow up on a sufficient number of evaluations to derive meaningful lessons to inform the work of the Board. However, the office should also be small enough to force the prioritization of topics and the co-ordination of its efforts with evaluation underway elsewhere in the Fund. In addition to detailed knowledge of the macroeconomic issues core to the Fund's mandate, the staff of EVO should embody broad experience and backgrounds to permit EVO to engage in assessments on a wide range of topics. Where additional experience or perspective is needed, EVO would be provided with a budget from which it could augment its staffing on an evaluation-by-evaluation basis with external consultants and experts to participate in, lead, or even wholly conduct particular evaluations. This would be one channel through which to address the desire of some to ensure that external input formed a part of independent evaluation. Where appropriate, external input could also be obtained through public consultations conducted by EVO, with assistance from EXR.

VII. NEXT STEPS

59. If the recommendation for Option 3 is accepted, staff—under the direction of the EG—would proceed to prepare a proposal on how to operationalize EVO for consideration by the EG and Executive Board. A Charter would need to be drafted for EVO and the Terms of Reference for the EG would need to be amended. The articulation of accountabilities would need to ensure the independence of the unit and its Director from Fund management

and their operational independence from Executive Board. Clear principles for the co-ordination of evaluation efforts throughout the Fund would also need to be articulated, including a clear understanding of the division of labour between EVO and OIA.

60. The number of new staff positions needed will depend on the expected size of the work program and on the size of the budget for engaging external consultants and experts. In addition to a Director for EVO, staffing in a modest range (both professional and administrative and research support) would be required. This need not be fully additional to baseline levels since consideration could be given to the possible transfer of a limited number of staff positions from a more-focussed OIA and/or other departments. However, the establishment of EVO would not likely be feasible without some increase in overall staffing.

61. Staff should undertake a formal assessment of resources requirements, including the scope to re-allocate resources from elsewhere within the Fund.

VIII. APPENDIX I: TERMS OF REFERENCE OF THE EVALUATION GROUP²¹

A small group of Executive Directors shall be designated by the Executive Board, to follow closely the evaluation function in the Fund, and advise the Executive Board.

The composition of the group will be proposed by the Chairman of the Executive Board, in consultation with the Dean, and approved by the Executive Board. It should normally be composed of four Executive Directors, representing a balance of interests. Periodic rotation of membership should occur, on a staggered basis, to enable different members of the Board to have an opportunity to be members, while ensuring a sufficient degree of continuity of involvement with each evaluation project. All members of the Board may, however, attend any meeting of the group and participate in its deliberations.

The group will consider proposals for evaluation topics emanating from the Board. Topics may include those that could be undertaken entirely within the institution (by the Executive Board or by the staff) and those that could be undertaken jointly by staff and outside experts, or those that could be undertaken entirely by outside experts. In the case of topics that would involve outside experts, the group would consider the choice of evaluation projects, their possible scope, the appropriate methodology, the choice of outside experts, whether the findings should be published, and other elements of proposed evaluation studies (including, for example, their budget and overall time frame). Based on the group's discussions, and after consultation with the Management, the chair would make recommendations on all these aspects to the Executive Board for its approval. Once an evaluation project is approved by the Board, the group would monitor its progress on a continuing basis. In the case of projects that would be carried out by the staff, the staff would consult with the group on the coverage and design of the project to ensure that it would address the concerns of the Executive Board.

It is envisaged that normally there will be no more than two or three external evaluations per year.

The experience with this method of conducting and monitoring the evaluation function in the Fund will be reviewed in early 1998.

²¹ EBD/96/102, Supplement 1 (9/9/96)

IX. APPENDIX II : EVALUATION STUDIES 1996-99

62. The annex reviews 12 of the evaluation studies conducted during 1996-99 . The source material comprises the views of Executive Directors, staff, and management regarding the technical merits, findings, recommendations, and other features of these studies as contained in the minutes of the Board discussions and staff and management's formal responses to the evaluation reports.

1. Self-evaluations

63. The self-evaluations carried out since 1996 included the 1997 review of surveillance,²² the review of the ESAF,²³ the lessons for surveillance from the Asian crisis,²⁴ and the review of Fund-supported programs in the context of the Asian crisis.²⁵ As noted by Executive Directors during the respective Board discussions, all reports were of high quality and made valuable suggestions for improving the operations of the Fund. On methodology, Directors commended the design of the review of the ESAF which used a combination of cross-section analysis and case studies; however, some Directors felt that the case studies yielded more insight than the cross-section analysis.

64. Some Directors expressed dissatisfaction with the tone of the reports, in particular, the entire report on lessons for surveillance from the Asian crisis and the Executive Summary of the review of programs in the Asian crisis. While many Directors commended the staff for the candor of these reports, a few Directors indicated that the staff could have been more critical of the Fund's performance. Moreover, a few Directors, in the discussion of the lessons for surveillance from the Asian crisis, observed that the report had not been sufficiently critical of the role of the Executive Board in not doing more to forestall the crisis.

65. There was a mixed record of publication of the self-evaluations. The biennial review of surveillance in 1997 and the lessons of surveillance from the Asian crisis were not published, but extensive summaries were contained, respectively, in the 1997 and 1998

²² *Biennial review of the Implementation of the Fund's Surveillance: SM/97/53 (2/19/97), SM/97/92 (4/10/97), and EBM/97/24 (3/14/97).*

²³ *Review of Experience Under ESAF-Supported Arrangements: EBS/97/112 (6/23/97), EBS/97/123 (7/2/97), and EBM/97/75 (7/21/97).*

²⁴ *Review of Members' Policies in the Context of Surveillance - Lessons for Surveillance from the Asian Crisis: EBS/98/44 (3/9/98), and EBM/98/34 (3/26/98).*

²⁵ *Fund-Supported Programs in the Asian Crisis: EBS/98/202 (11/25/98), and EBM/98/130 (12/18/98), and Summing-Up by the Chairman (revised): BUFF/98/117.*

Annual Reports of the Fund.²⁶ Reflecting the desire of the Executive Board for greater transparency about Fund activities, the report on the review of the ESAF and the review of programs in the Asian crisis were published in full, together with the Summing Up of the respective Executive Board discussions.

2. External evaluations

66. This section focuses on the experience with the three external evaluations decided upon by the Executive Board and carried out under the guidance of the EG. In addition, there were two external evaluations ongoing or just completed at the time the present evaluation structure was initiated in June 1996 and four external evaluations of aspects of the Fund's external communications carried out in 1999.²⁷ The 1999 studies provided the background information for a strategy to strengthen the Fund's external communications over the medium term²⁸ but are too specialized to be reviewed in this paper. However, lessons for the evaluation function may be derived from the two earlier studies. These were the evaluation of the technical assistance provided by MAE, and the evaluation of three aspects of program design in European transition countries.

(i) External evaluation of technical assistance provided by the Monetary and Exchange Affairs Department²⁹

67. For this study, the three-member expert panel selected by MAE determined its own approach to fulfilling the terms of reference, choosing to examine technical assistance (TA) in 20 countries out of some 130 countries that had received TA in 1991-94. The direct cost of the study (evaluators' fees and expenses) was \$258,000 and the indirect cost – staff time spent by MAE in supporting the study – was 0.4 staff years. Because the study had to be completed in about six months (the second half of 1995), it was not possible to study the

²⁶ Although the Executive Board endorsed full publication of the report on lessons for surveillance from the Asian crisis, management decided not to proceed due to concerns related to legal proceedings in a member country.

²⁷ The four communications studies were: "Moving the IMF Forward: A Plan for Improving the Fund's Communications with Critical Audiences Around the Globe" by Edelman Public Relations Worldwide, June 1999 (which incorporated results from surveys by Wirthlin Worldwide); "The Effectiveness of IMF Communications: View from the Markets" by Susumu Awanohara, March 1999 (with a follow-up assessment in August 1999); "The IMF and the U.S. Congress: An Uphill Struggle" by Mary Locke, May 1999; and "Assessment of IMF Media Operations During the Interim Committee Meetings" by James Morgan, April 1999 (with a follow-up assessment in October 1999).

²⁸ *Strengthening the Fund's External Communications – Plans and Resource Implications*; SM/00/14 (1/27/00).

²⁹ *External evaluation of technical assistance provided by the Monetary and Exchange Affairs Department – Report of Independent Panel*: EBS/96/15 (1/26/96) and EBM/96/47 (5/17/96).

results of many of these projects when their outcomes had reached full development. In the Executive Board discussion of the report (in May 1996), Directors generally endorsed the need to improve the monitoring and evaluation of TA activities, and urged the staff to enhance its practice of self-evaluation. However, no formal action plan to implement the recommendations of the report was shared with the Executive Board, and some of the review's findings (including the need to introduce formal self-evaluation of projects) were repeated in the 1998 internal review (see below). Some Directors also suggested that a more in-depth study of a smaller sample may have yielded different findings to those of the review. They proposed that for the future, a body of internal ex-post evaluations should be built up over time, which could be used as the foundation for an external review. Many Directors also requested to be consulted in future when external reviews were initiated; this issue was addressed in the Board's endorsement of the evaluation structure in June 1996 which included the establishment of the Evaluation Group of Directors. The evaluation report was not published, but a summary of the panel's findings and the Board response was included in the Fund's 1997 Annual Report.

(ii) External review of aspects of program design in European transition economies

68. This evaluation consisted of studies by outside experts of medium- and long-term aspects of fiscal performance, monetary policy, and exchange rate policy confronting transition economies. The direct cost of the study (evaluators' fees and expenses) was \$65,000, and the indirect cost – staff time spent by EUI and EUII in support of the study – is estimated at 0.4 staff years. The original intention of the sponsoring departments—EUI and EUII—was that the findings should be discussed by the Executive Board as a guide to future practice. However, the quality of one of the studies was mixed, so that management determined that the papers should not be proposed for discussion by the Board. Instead the papers were published as Fund Working Papers (WP/96/108, WP/96/125, and WP/97/31).

(iii) External Evaluation of the ESAF³⁰

69. This was the first of three studies carried out under the guidance of the Evaluation Group of Executive Directors. For each of these evaluations, the report itself, the staff response, the statement by the Chairman of the EG at the Executive Board discussion, and the Summing Up of the Board discussion were published to enhance transparency. For the last two evaluations (concerning surveillance, and research activities) a management statement was also published.

³⁰ *External Evaluation of the Enhanced Structural Adjustment Facility: EBAP/98/8 (1/22/98); Staff Response to External Evaluation of the ESAF: EBS/98/33 (3/2/98); and EBM/98/25 (3/11/98). The Summing Up of the Board discussion, the evaluation report, the staff response, and the statement at the Board discussion of the Chairman of the EG were published together for a press conference on March 13, 1998.*

70. Many Directors commended the design of the external ESAF evaluation in that it complemented that of the internal review of the ESAF being carried out by staff at the same time (see above). The three topics chosen as the focus of the evaluation were those on which either a fresh perspective was needed, where there was a need to build credibility outside the institution, and/or where the Fund was deemed to lack adequate expertise. They were: (i) the development of countries' external positions during ESAF-supported programs, i.e., how to gauge progress toward external viability; (ii) social policies and the composition of government expenditure, i.e., how to integrate social considerations into a macroeconomic program; and (iii) the determinants and influence of differences in the national ownership of programs.

71. These questions were to be addressed by in-depth case studies; the terms of reference specified that the evaluators were to select the country cases, subject to guidelines of (a) 4-7 countries per topic with maximum overlap over the three topics, (b) ensuring geographical diversity, and (c) including both strong- and weak- performing countries. The terms of reference also specified that the cost of the study was to be limited to \$600,000 and that it was to be completed in about nine months between the first quarter of 1997 and the fall of 1997. In the event, direct costs (evaluators' fees and expenses) amounted to under \$559,000, but there was also the indirect cost of the time spent by OIA staff in logistics support for the evaluation (including overtime) of 0.9 staff years, not counting time spent by staff in area and functional departments in assembling information for the evaluators.

72. The evaluators interviewed a wide cross-section of stake-holders in six countries with regard to external developments, in five countries concerning social policies, and in seven countries on the topic of national ownership. They also had access to all relevant Fund documents and used other written sources in their work.

73. While Directors agreed with the key elements of the report's recommendations—the need for greater cooperation with the World Bank on social impact and other macroeconomic policies, and the steps to be taken by the authorities and the Fund in order to promote ownership of programs—some of the recommendations and the findings on which they were based were not endorsed. Moreover, some Directors believed that the evaluators did not fulfil the terms of reference: “with regard to external viability, the report did not systematically analyze the reasons for the diverging experiences of the countries in the sample.”

74. Some Directors also considered that the quality of the evaluation was impaired to some degree by the lack of transparency in the methodology used by the evaluators. Directors acknowledged that the small number of case studies allowed by the terms of reference could lead to findings that were peculiar to individual cases and possibly contradictory to each other. But it was difficult in some instances to determine which of the study's findings were based on empirical research, which on the case studies, and which on *a priori* reasoning or

normative analysis. There were some unsubstantiated assertions and recommendations, which, even if correct and reasonable, could not be seen to have been empirically based.

(iv) External evaluation of Fund surveillance³¹

75. A new feature of this study was the hiring of a full-time research associate to assist the evaluators, which permitted a greater depth of coverage of cases and also reduced somewhat the logistical support required of OIA. In addition, management, for the first time, prepared a statement of its views on the evaluation report and the staff response for the consideration of Executive Directors.

76. The study was regarded by Executive Directors as being of very high quality. As noted in the Summing Up of the Board discussion: "Executive Directors welcomed the report They expressed their deep appreciation for the careful work and considered judgements of the panel. Directors considered that the issues raised in the report would serve to stimulate debate within and outside the institution, and to motivate further discussion of a number of topics of importance to the work of the Fund."

77. The evaluators performed in-depth studies of Fund surveillance in 12 countries over a ten-year period, using all documentary information available in the Fund, and interviewing a number of officials and other observers in each country who were knowledgeable about Fund surveillance. In addition, interviews were held with about 50 Fund staff members, a majority of Executive Directors, a number of officials of international institutions and the EU, 11 academics, 5 representatives of NGOs, and 31 representatives of the private sector. Although the staff response noted the non-random method of selection of interviewees, this criticism was not endorsed by Directors who regarded the selection of countries and interviewees as broad and representative.

78. In light of experience with the external evaluation of the ESAF, the budget limit was raised to \$700,000, and the time needed to carry it out was increased to 11 months (July 1998 – June 1999). In the event, the evaluators' fees and expenses were below the limit at \$689,000, but a still considerable time (0.8 staff years including overtime) was required from OIA staff to support the study, not counting time spent by staff in area and functional departments in assembling information for the evaluators.

79. Some 20 of the 29 specific recommendations were endorsed by most Directors.

³¹ *External Evaluation of Fund Surveillance: EBAP/99/86 (7/15/99); Staff Response to the External Evaluation of Fund Surveillance: EBAP/99/88, Rev.1 (9/7/99); Statement by the Managing Director on the Report of External Evaluators on Fund Surveillance—Further Consideration: BUFF/99/102 (8/24/99); and EBM/99/94 (8/27/99). The statement at the Board discussion by the Chairman of the EG, the Summing Up of the Board discussion, the evaluation report, the management statement, and the staff response, were published together on September 13, 1999.*

80. There were a number of “evaluation function” issues that emerged from this study, including the need to institute and monitor an effective mechanism to ensure that the recommendations endorsed by the Executive Board are incorporated into the surveillance work of the Fund, and to undertake more systematic evaluations of surveillance. With regard to the feedback mechanism, a number of Directors agreed with the panel that the recommendations from past internal reviews had not been incorporated into ongoing operations.³² It was further agreed that after a period of 12 months, the Board should review how well the action plan had been implemented.

81. With regard to introducing more systematic evaluation, the evaluators recommended that the Fund experiment with *ongoing* external reviews of a sample of Article IV staff reports (#25). This recommendation was endorsed by Directors, and has been implemented informally by inviting outside comment on Article IV staff reports published on the Fund’s Internet site.

(v) External evaluation of the Fund’s research activities³³

82. As with the external evaluation of surveillance, this study included the hiring of a full-time research associate by the evaluators and the issuance of a management statement at the request of the Executive Board. As indicated in the Summing Up of the Board discussion, the study was highly regarded by Executive Directors, who “considered that the evaluators had done a valuable job in judging whether the Fund’s diverse economic research output met the multiple expectations placed on it.” However, many Directors also agreed with the staff response that the methodology used was not sufficiently comprehensive to arrive at findings which could be accepted without further careful consideration. They noted that the time and resource constraints contained in the terms of reference—the study was to be completed in seven months at a cost of not more than \$220,000—did not allow the evaluators to consider the full range of research activities in the Fund or to interview a broadly representative sample of the users of Fund research (African and Middle Eastern countries were significantly underrepresented).³⁴ Moreover, the selection of research output to be judged for

³² In the press conference following the publication of the report on September 14, 1999, the Chairman of the panel noted that “the lessons of Mexico had not been well absorbed, for the most part – one clear criticism we make”.

³³ *External Evaluation of the Fund’s Economic Research Activities: EBAP/99/85 (7/15/99); Staff response to the External Evaluation of the IMF’s Research Activities: EBAP/99/87 (7/15/99); Comments by the Managing Director on the Report of External Evaluators on the Fund’s Economic research Activities—Further Consideration: BUFF/99/109 (9/2/99); and EBM/99/99, (9/7/99). The statement at the Board discussion by the Chairman of the EG, the Summing Up of the Board discussion, the evaluation report, the management statement, and the staff response, were published together on September 14, 1999.*

³⁴ *The direct cost of the study was \$246,000 and the indirect cost (support by OIA staff) was 0.4 staff years, not counting time spent by staff in area and functional departments in assembling information for the evaluators.*

quality was limited to that in one year (1998) when a longer-term perspective (perhaps over a five-year period) would have been preferable. Finally, the diversity of the Fund's research output might have required the panel of evaluators to be enlarged to include experts in areas outside the three evaluators' specialties.

83. Twenty of the 22 recommendations were either supported by Directors or would be considered for implementation as part of broader reviews of Fund procedures.

84. The issues related to the Fund's evaluation function that emerged from this study also involve an improved feedback mechanism and more systematic evaluation. With regard to feedback, Directors endorsed both the suggestion by the Chairman of the EG that management prepare an action plan to incorporate the recommendations of the report into the Fund's research activities, and the evaluators' recommendation that the Board should review progress on implementing the recommendations of the study after one year's experience (#21).

85. With regard to more systematic evaluation, Directors endorsed the recommendation to conduct periodic, general, external evaluations of research (#22). These reviews could be part of the Fund's future evaluation program under the oversight of the EG. Directors also endorsed the recommendation that, as in the World Bank, outside experts should review a selection of research projects each year (#20) to make it more likely that Fund research will incorporate the latest ideas and developments.

3. Internal Evaluations by the Office of Internal Audit and Inspection (OIA)

(i) Review of the resident representative program³⁵

86. In mid-1996, management, faced with pressure to increase the number of resident representative posts on the one hand and find budgetary savings on the other, initiated a full-scale evaluation of the effective use of Fund resident representatives. OIA had previously, in late 1995 and early 1996, reviewed the adherence by resident representatives to the guidelines for handling administrative arrangements at posts.

87. The Summing Up of the Board discussion noted that the study was "candid, balanced, and comprehensive". This assessment reflected in part the review's methodology which was notable for its transparency. All departments concerned with resident representatives were involved in the design of the review and were informed of its findings as the review proceeded. Interviews were held with Executive Directors, staff, and national authorities concerning the nature and scope of the review. These inputs, and the assistance of a survey

³⁵ *Review of the Resident Representative Program: EBS/97/137 (7/25/97) and Supplement 1 (9/17/97); and EBM/98/7 (1/23/98). This review was not published but a short summary was included in the 1998 Annual Report of the Fund.*

design consultant, were used to design a survey with questionnaires tailored to different respondent groups.

88. In December 1996, the survey was sent to 600 respondents in four groups comprising Central Bank Governors and Ministers of Finance; current and former resident representatives; staff mission teams; and Executive Directors, Fund management, and senior staff. In March-May 1997, OIA staff visited a representative sample of five countries for follow-up discussions of tentative conclusions and possible changes to the resident representatives program. Similar discussions were also held in Washington with officials from a number of countries. Finally, prior to the finalization of the report, OIA staff held round table discussions with representatives of all concerned Fund departments on the merits and shortcomings of possible alternative reforms.

89. Six OIA staff were assigned to the project for various periods over the total project span of 17 months. The overall resource cost was 3.4 staff years of regular time and 0.6 staff years of unpaid overtime. In addition, dollar costs of \$27,000 were incurred for staff travel and \$61,000 for the fees of two consultants.

90. Of the 33 recommendations relating to the role of the resident representative, the personnel selection process and program administration, partnership arrangements with member countries, and targeting the program to member countries, 24 were endorsed by most Directors. Directors also endorsed the nine recommendations relating to the budgetary framework of the program. Most of the recommendations endorsed by the Board have now been implemented. In addition to demonstrating that an independent internal review can be effective if it is well designed and carried out, this study again evidenced the need for a systematic approach to evaluation in the Fund. The Board agreed that a focussed review of the resident representatives program (to examine how many of the recommendations endorsed by the Board had been implemented) should be held after two years (in FY 2000), and a full review of the effectiveness and efficiency of the program should be held every five years.

(ii) Review of Fund Technical Assistance³⁶

91. All departments concerned with technical assistance (TA) in the Fund were consulted on the issues to be examined in the review, the evaluation methodology, and the design of surveys. Against this background OIA, in consultation with the EG, decided to undertake (i) an extensive review of available documents (TA procedures and reports); (ii) two surveys on technical assistance provided by FAD, MAE, and STA and a separate review by an outside

³⁶ *Review of Fund Technical Assistance: EBAP/99/59 and Supplement 1 (5/17/99); staff response – Statement on the Review of Fund Technical Assistance and Suggested Issues for Discussion: EBAP/99/60 (5/17/99); and EBM/99/61 (6/8/99). These documents were not published either in full or in summary form.*

consultant of TA provided by the Bureau of Computing Services (BCS) during 1996-98; (iii) interviews in the field and at headquarters with country authorities; discussions were held with officials from 19 countries in late 1998; and (iv) discussions with staff focus groups.

92. The first survey sought the general views concerning TA of over 1,000 potential respondents – Fund staff, Executive Directors and their Alternates, long-term experts, and officials of countries that had received substantial technical assistance in the recent past. The second survey was an impact evaluation study of 100 randomly selected projects out of the 997 projects implemented in FY 1996-97. All impact surveys were completed by Fund area and technical assistance departments, but only 46 impact surveys were completed by the authorities in the recipient countries.

93. Seven OIA staff were assigned to the project for various periods over the total project span of 19 months, for a total of 5.4 staff years of regular time and 0.9 staff years of unpaid overtime. In addition, dollar costs of \$45,500 were incurred for staff travel and \$151,000 for the fees of four consultants.

94. Executive Directors regarded the report as being comprehensive and of high quality and accepted most of OIA's recommendations. The response of the staff involved in the administration and implementation of TA, however, advocated a more cautious approach, due to concerns about the cost of their implementation and reservations about the methodology used in the review. Consequently, while the broad thrust of the study's recommendations was endorsed by the Board, it was also agreed that some of the more detailed recommendations should be subject to further study or should be introduced on an experimental basis because of the possible additional resource costs.

95. With respect to follow-up, inter-departmental working groups have prepared for Board consideration a policy statement on technical assistance; a policy paper on country contributions for technical assistance services, and a description of the evaluation methodologies of other TA providers. In addition, an Annual Report on TA activities is being prepared for Board discussion in August 2000 which is expected to include proposals for operational guidelines for improvements to TA activities arising from the study.

(iii) The General Services Review

96. This study comprised a two-year comprehensive assessment of the effectiveness and efficiency of support services in the Fund comprising thirteen different activities in the areas of information services, facilities and related services; and financial support and control services. It used a consistent methodology, validated by an independent consultant, to evaluate each activity. This was to identify and analyze the inputs involved in producing the service; examine the volume of the outputs and their dollar and system costs; assess the quality of the outputs using available reviews, previous surveys, and new surveys; and compare the results with "best practice" benchmarks. Transparency was optimized by seeking the views of providers and users at every stage, and by providing departments with

the opportunity to comment on the draft report. Any difference of views which remained were reflected in the final report.

97. Twelve OIA staff were assigned to the project for various periods over the total project span of two years, for a total of 11.3 staff years of regular time and 1.1 staff years of unpaid overtime. In addition, dollar costs of \$703,000 were incurred for the fees of three consultants.

98. The study began to have an impact on Fund operations well before its completion. Action plans to remedy identified deficiencies were proposed at each stage of the review, and these are being implemented by the departments concerned. One such example is the reorganization and associated streamlining of the Administration Department, the Bureau of Computer Services, the Bureau of Language Services, and parts of the Secretary's Department into the Human Resources Department and the Technology and General Services Department.

99. A summary of this study was provided to the Executive Board for information and was not discussed as a separate agenda item.³⁷ However, some Executive Directors commented on the review during the discussion of the Administrative and Capital Budgets for FY 2000 (EBM/99/45, 4/20/99). In this discussion, many Directors noted the substantial increase in the staff ceiling for FY 2000, and called for an external review of the Fund's operating processes and procedures. In this context, some Directors commended the quality of the General Services Review and indicated that it should serve as a foundation for the external review. As to the review itself, one Director characterized the main recommendations as "compelling", and urged that these recommendations continue to be implemented in the areas where they had begun, and that they soon be implemented in other areas.

³⁷ *General Services Review, EBD/99/52 (4/9/99). A short summary of this document was published in the 1999 Annual Report of the Fund.*