

**ANALYSIS OF KEY ISSUES**  
**EMERGING FROM THE AFRICA BUREAU**  
**FY 1999 R4 ASSESSMENT**

**November 3, 1997**

Prepared by:  
Anne Inserra and Janet Tuthill of Management Systems International  
for the Africa Bureau, U.S. Agency for International Development  
under a task order managed by PPC/CDIE, under the  
Performance Measurement and Evaluation IQC.

# TABLE OF CONTENTS

## FOREWORDii

### I. INTRODUCTION 1

#### A.BACKGROUND OF THE PAPER1

#### B.METHODOLOGY USED IN THE ASSESSMENT2

### II.ANALYSIS OF KEY ISSUES IDENTIFIED IN THE ASSESSMENT & RECOMMENDATIONS FOR ADDRESSING THEM4

#### A.WERE THE OBJECTIVES FOR THE FY 1999 R4 MET?4

##### 1.Agency-wide R4 Guidance4

##### a.Clarity and adequacy of Agency-wide R4 guidance4

##### b.Timeliness of Agency-wide R4 guidance7

##### c.R4 page limit7

##### 2.Africa Bureau Supplementary Information on the R410

##### 3.Success of the R4 in Meeting its Objectives10

##### a.Clear understanding of the objectives for the R411

##### b.Success of the R4 in terms of the objective of assessing progress toward results12

##### c.Success of the R4 in terms of the objective of using performance information to influence budget allocation decisions15

##### d.Managing for Results and the R417

##### e.The feasibility of meeting all of the purposes for the R418

#### B.HOW WELL DID THE AFRICA BUREAU R4 PROCESS WORK?19

##### 1.Budget Categories in the Performance Scoring Process20

##### 2.Complexity of the Review Process21

##### a.Performance scoring process and criteria21

##### 3.Feedback and Communication between Washington and Operating Units23

##### 4.Technical Review Committees24

##### 5.Newer Strategic Objectives25

##### 6.Program Review Days26

##### 7.Senior Management Participation in the R4 Review Process26

##### 8.Efficiency of the R4 Preparation and Review Process - Was the Level of Effort Appropriate?26

##### 9.Recommendations on How to Improve the Effectiveness and Efficiency of the Africa Bureau R4 Process28

#### APPENDIX A:Lists of Respondents and Persons Interviewed

#### APPENDIX B: Interview Guide for Key Informant Interviews and Focus Group Interviews

#### APPENDIX C: Questionnaire for Operating Unit Survey

## **FOREWORD**

This paper presents analysis of some of the key issues that emerged from the Africa Bureau's assessment of the FY 1999 R4. It is a companion piece to the document entitled "Briefing on Preliminary Findings & Recommendations from FY 1999 R4 Assessment," that was presented by MSI to the Africa Bureau on October 15, 1997. This paper focusses on more in-depth analysis and explanation of some of the most important issues and recommendations that were identified during the course of the assessment and highlighted in the "Briefing" paper. Readers are encouraged to consult both documents for a more complete understanding of the key assessment results.

# **I. INTRODUCTION**

## **A. BACKGROUND OF THE PAPER**

Upon completion of the FY 1999 R4 review process, the Africa Bureau determined that it would like to assess the effectiveness of the entire R4 effort for the previous year. Assistance from Management Systems International to conduct the assessment was contracted through the Africa Bureau buy-in to a task order managed by PPC/CDIE, (under the Performance Measurement and Evaluation IQC). A working committee with representatives from the Africa Bureau's Development Planning (DP) and Sustainable Development (SD) offices met several times to plan the scope and specifics of the assessment during July of 1997. Representatives from other Africa Bureau offices, M Bureau, and CDIE were also consulted. The data collection effort for the assessment got underway in August, and was completed in September.

MSI staff presented preliminary findings and recommendations emerging from the assessment data at a special briefing on October 15, 1997 held for Africa Bureau (including senior staff), with representatives from G, M, and PPC Bureaus also invited. The document prepared for that meeting is a companion piece to this assessment report, and is entitled "Briefing on Preliminary Findings & Recommendations from FY 1999 R4 Assessment." This paper focusses on more in-depth analysis and explanation of some of the most important issues and recommendations that were identified during the course of the assessment and highlighted in the "Briefing" paper. Readers are encouraged to consult both documents for a more complete understanding of the key assessment results.

Both the analysis and briefing papers represent attempts to accurately portray the perspectives of those who were consulted in the assessment in both the presentation of findings and the framing of recommendations for the future. As is described below, the assessment gathered mostly qualitative information from individuals who "self-selected" to respond. Because of this, the assessment data do not lend themselves to detailed statistical analysis. The authors were encouraged by the Africa Bureau sponsors to use their own judgement to some degree in attempting to synthesize the enormous amount of information gathered, which also lent an element of subjectivity to the analysis. The authors were also requested to draw on their collective experience to "digest" the findings and hone in on key recommendations for the future.

Guiding principles for the analysis process were to identify the most commonly held views where possible, and to identify differences in perspective between groups (such as Washington vs. the field) if obvious. It is interesting to note that despite the richness and diversity of information collected, consensus did emerge from the data around a set of major findings and recommendations. This consensus around some key issues seemed to be further supported by the reactions of those who attended the briefing presentation on preliminary findings and recommendations that was held with Africa Bureau on October 15.

The Africa Bureau plans to use the assessment findings and recommendations as the starting point for discussions on how to design the R4 preparation and review process for next year. Africa Bureau plans to share its findings widely within the Bureau and with the rest of the Agency. Through a consultative process involving an Africa Bureau working group, cooperation with the Inter-Agency task force that is developing the Agency-wide R4 guidance, and discussion at the Africa Bureau Mission Directors' conference, it is planned that a set of recommendations will emerge to guide planning for the Africa Bureau FY 2000 R4 process.

## **B. METHODOLOGY USED IN THE ASSESSMENT**

The purposes for the R4 assessment included the following:

- to determine whether the FY 1999 R4 met its objectives;
- to assess the effectiveness of the Africa Bureau FY 1999 R4 review process;
- to assess the efficiency of the Africa Bureau FY 1999 R4 review process; and
- to gather suggestions on how to improve the R4 preparation and review process for the following year.

The Africa Bureau working committee determined that the assessment methodology would be highly participatory. Virtually all Africa Bureau staff who participated in the review process and/or in the preparation of a FY 1999 R4 were offered an opportunity to provide input. In addition, key representatives from G, M, and PPC (CDIE) Bureaus who had participated in some way in the Africa Bureau R4 review process were invited to participate in interviews.

Another major principle for the assessment was that direct attribution of comments and opinions to specific individuals would not be made in the presentation of findings, in an attempt to encourage an atmosphere in which comments might be offered more freely. (It is interesting to note that in the end the majority of respondents to the operating unit questionnaire did elect to include their names on the response sheets, even though doing so was optional.)

Three basic techniques of a questionnaire survey of operating units, focus group interviews, and key informant interviews were used to collect information on views on the R4. Data collection instruments were designed by MSI staff, and approved by the Africa Bureau working group. Information was collected by a team of MSI staff, including Janet Tuthill, Anne Inserra, David Hirschmann, Drew Lent, Judy Light, and Donald Spears. Analysis of the collected data was conducted by Anne Inserra and Janet Tuthill, assisted by Hope Davis.

The first technique used in the assessment was a questionnaire survey for operating units. The questionnaire was sent to all operating units that had prepared a FY 1999 R4,

including missions, regional offices, and Washington-based units. Any number of individuals within an operating unit were invited to complete the questionnaire. Respondents were permitted to remain anonymous, although most chose to identify themselves. Forty-seven responses were received. In some cases, more than one individual from an operating unit responded. In a few cases, units sent in a combined response representing the responses of several individuals within the unit. Representatives of an identified seventeen different operating units responded.

The second major technique used in the assessment was that of focus group interviews. Africa Bureau country desk officers (CDOs) and virtually every individual who had participated in an Africa Bureau R4 technical review committee (these committees reviewed the R4s and determined the performance scores for strategic objectives) were invited to participate in focus groups. A total of seventy-eight people were invited, and thirty-three accepted and attended one of eight focus group sessions. There was an attempt to have CDOs and technical review committee members in separate focus groups, as well as to have geographic diversity amongst the offices represented within each focus group, although these principles sometimes broke down due to scheduling practicalities. An interview guide was used to prompt the sessions, which generally lasted anywhere from approximately an hour and a half to slightly over two hours in length. A moderator guided discussion, while a recorder took notes. Africa Bureau offices represented among the attendees included EA, SA, WA, SD (DRC, HRD, PSGE, and SA), and DP (POSE). Other Bureaus represented included G (DG, ENV/EET, and PHN), M (B/PA), and PPC (CDIE/RRS and PHD).

The third technique involved key informant interviews. Twenty-three individuals from Africa, PPC, G and M Bureaus were invited to have an individual key informant interview. In addition, desk officers were asked to identify any representatives from field operating units who might be in Washington during the relevant time period and who might wish to be interviewed. Fifteen persons responded to the invitation and received an interview. Those interviewed represented different Washington-based offices within Africa Bureau (DP, EA, SA, WA, and SD), plus one person from CDIE. An interview guide was used for all of the interviews. Interviews ranged from approximately thirty to ninety minutes in length.

It was stressed during the design that the effort would be an assessment, not a formal evaluation. It was recognized from the start that most of the data to be collected would be imprecise and qualitative in nature, and represent subjective opinions of respondents. All of the information collected through the key informant and focus group interviews was free response. Most of the data collected through the questionnaire were also free response, although a number of scaled response and yes/no questions were included. (See Appendix C, which is the actual questionnaire instrument used for the operating unit survey. Question # 4a is an example of a scaled response question; question # 4c is an example of a yes/no question.) In addition, no rigorous sampling procedures were used. Instead virtually everyone who either prepared a FY 1999 R4 or participated in the Africa Bureau review of R4s was invited to participate and individuals decided whether or not to respond or participate in an interview or focus group. This resulted in a data set that represented the views of a

group of self-selected respondents. The implication of all of these factors was that the resulting data did not permit sophisticated statistical analysis, and findings were not statistically valid. These facts were anticipated during the design of the assessment, and felt to be appropriate for the nature of the analysis needed.

Through the use of the three major methods described above, information was gathered from over 100 individuals representing a wide range of offices, including some offices outside the Africa Bureau. This tremendous response indicates the great interest in the R4 assessment. The information collected was nearly overwhelming in its richness, complexity, diversity, and sheer quantity. The MSI team was extremely impressed by the thoughtfulness of comments provided, as well as the willingness of so many staff to take time to participate in the assessment.

## **II. ANALYSIS OF KEY ISSUES IDENTIFIED IN THE ASSESSMENT & RECOMMENDATIONS FOR ADDRESSING THEM**

### **A. WERE THE OBJECTIVES FOR THE FY 1999 R4 MET?**

One of the first issues to surface during the preliminary design of the assessment was that there exist a number of different purposes or objectives for the Results Review and Resource Request (R4). The Automated Directives System (ADS) lays out a number of different purposes for the R4, and the Agency-wide guidance (STATE 002636) for the FY 1999 R4 described some implicit purposes which were in basic harmony with the ADS but did not exactly correspond to it. In addition to differences that seem to appear in these official sources of guidance, there are different phases or parts of the R4, including preparation and review of the R4 internally within the operating unit, and the USAID/Washington-based review of the R4. The result of all this is that there is a multiplicity of different views and interpretations of the purposes for the R4 among Agency staff. Some of these differences in view may be linked to the organizational affiliation of individuals, whether in terms of belonging to a specific Bureau, or to a field operating unit as opposed to a Washington-based office. Some differences may be linked to the direct experience that an individual staff person has in the various phases or uses of the R4, e.g. preparing the document, or reviewing it. In recognition of this complexity, the assessment tools followed up the general question of whether staff felt that the objectives had been met with more specific probes on different individual purposes. The information collected confirmed the initial assumption that there are many different interpretations of the purposes for the R4.

#### **1. Agency-wide R4 Guidance**

##### **a. Clarity and adequacy of Agency-wide R4 guidance**

Staff were asked about the usefulness of the Agency-wide FY 1999 R4 guidance, as laid out in STATE 002636. The great majority of operating unit survey respondents found the guidance to be “moderately clear” or better, and “somewhat adequate” or better. A positive comment was that it provided fairly clear instructions and was an attempt to make the process less burdensome. Feedback from the key informant and focus group interviews showed that those who thought the guidance was fine were in the minority. Some people found the guidance to be too lengthy, complicated, and/or difficult to read. A number of people stated that the guidance kept changing or evolving, and that different documents were inconsistent and contradictory. They felt this was confusing, and caused some operating units to do revisions of their R4s several times.

One of the most important aspects of the Agency-wide R4 guidance that was not clear to everyone, as shown by several survey and interview comments, was which indicators (for which intermediate results), and how many indicators (how many in total, and how many for specific results such as the SO) should be reported on in the R4. It became clear from the assessment findings that there is considerable disagreement about what information should be reported on in the R4. The ADS basically requires operating units to report in the R4 on progress toward achieving the SO, including data on intermediate results where appropriate, which in effect gives units leeway in terms of determining which IRs they will report on in any given year. The Agency-wide R4 guidance cable for FY 1999 seemed to match this intent, by instructing that “performance analysis should focus on the objective level and most significant ... IRs, i.e., those whose performance ... has been most important in determining the extent of progress toward achieving the objective.” The R4 guidance cable appeared to go even further than the ADS, however, in calling for “reporting by exception,” or emphasizing significant variations from plans. Assessment findings show that some people apparently feel that the requirement should be instead that all SO and IR indicators are reported on to Washington every year, while others think that selected IR indicators can be reported on, as long as the same ones are reported every year. Another element of confusion in the R4 guidance cable concerned the instructions for performance indicator tables, which stated that “each indicator used to monitor performance of the objective and/or key intermediate results” should be included. It was not clear whether this meant that the entire set of indicators for the SO and each of the IRs that are highlighted in the document should be included, or whether fewer, selected indicators for the SO and the key IRs could be used.

It is clear that decisions need to be made regarding the reporting requirements in the R4 so that expectations can be clearly laid out for next year. In pursuing this discussion, several facts seem particularly relevant. The first is that the assumption behind the ADS was that the full set of indicators for operating units, complete with baselines, targets, and actual data would be available to USAID/Washington at R4 review time through a working NMS system, which has not yet occurred. Secondly, the newer Agency requirement as stated in the FY 1999 R4 guidance cable that Bureaus must “cluster their (operating) units' objectives into three categories for funding purposes, initially on the basis of performance, then Agency goals and Bureau priorities,... within each of the Agency's strategic goal areas” was imposed after the ADS was written. There are indications that selective reporting in R4s as required by the

ADS (and even more so reporting by exception as stated in the R4 guidance cable), may not provide the level of detail on performance needed to be able to fulfill the newer Agency-wide requirement for clustering objectives. Africa Bureau technical review committees apparently thought that there were not enough performance data in many of the R4s to conduct the performance reviews, as evidenced by the high incidence of requests from the committees to operating units for further information.

A further point is that the R4 review process designed by Africa Bureau appears to have gone beyond the minimal requirements of enabling them to cluster objectives into the top twenty-five, middle fifty, and bottom twenty-five percent in each Agency goal area in several respects. In the Africa Bureau system, objectives were reviewed and clustered within even narrower budget categories than the Agency goal areas, a sophisticated ten-point scoring scale was used for each objective, and objectives were individually rank ordered within each category, instead of just clustered. Again, technical review committees often asked operating units for additional information beyond what was presented in the R4s in order to complete the scoring. It seems that the more precise and complicated review process opted for by the Africa Bureau may have created even greater demands for more detailed performance data in the R4 than either the ADS or the Agency R4 guidance cable asked operating units to provide. (Note: the ten-point scale system was intended to serve as a means to enable SOs to be clustered into the simpler four quartile system. It was developed in response to the difficulty of determining which quartile borderline SOs fell in. Some associated drawbacks of the more complex ten-point scoring system only became apparent later. See later discussion on the performance scoring process and criteria in section II.B. of this paper.)

The practical fallout of the above issues was that there was confusion both among operating units and technical review committees about which indicators and how many should be reported on in the R4. Operating units commented that in some cases when they did not include all indicators in the R4 document, technical review groups later asked for this information. Some felt they were penalized in the reviews for not including all the information. An operating unit staff person suggested that an annex be attached to the R4 that would contain the complete set of indicators and performance data for the program, to help reviewers who are not completely familiar with the program.

**Recommendation:** The following recommendation is proposed to address the confusion concerning what should be reported on in the R4 document.

- The R4 guidance should include clear instructions on how much and what type of information should be presented in the R4 document. More specific guidelines are needed concerning which indicators and how many should be reported. The principle of “reporting by exception” should be reconsidered.

Other comments on the Agency-wide guidance called for more details on how the scoring would be done, with clearer criteria for ranking categories. In general, there seemed to be a call for the guidance to be clearer, and less ambiguous. There was a sense that

operating units need to completely understand the criteria that will be used to judge the performance of their programs before they prepare their R4s.

**b. Timeliness of Agency-wide R4 guidance**

A slim majority of the operating unit survey respondents who had an opinion found the Agency-wide guidance cable to be timely. One person felt that receiving the guidance in early January provided plenty of time for operating units to prepare, discuss, review and approve the R4 in time for the late March submission deadline. A very significant proportion of survey respondents, however, felt that the Agency-wide guidance was not timely. A few people commented that the formats or charts were late in being sent out (presumably they were referring to the data tables). A number of comments from the operating unit survey responses called for the guidance to be sent out earlier, with a few specific suggestions for this to happen in November or early December. Several mentioned that receiving the draft guidance in November/December was very helpful to them. Reasons given for needing the guidance earlier included that R4 preparation requires so much time and effort that units need more time to rationally allocate resources to accomplish it, processes of internal review and gathering input from partners also take time, and some missions were nearly finished with R4 preparation by the time guidance arrived.

**Recommendation:** The following recommendation is proposed to address the timing issue.

- The R4 guidance should be sent out further in advance of the due date for the R4 documents. It should be sent at the end of November or early December for a March R4 due date.

**c. R4 page limit**

Many Washington- and field-based staff felt strongly that the R4 page limit of 30 pages as prescribed in the Agency-wide guidance cable was problematic. Some people seemed to think the specified limit was reasonable, and necessary for the review process, given the great number of operating units in the Africa Bureau that need to be reviewed every year. However, many objected strongly to the limit.

One problem was caused by the fact that some operating units adhered to the limit, while others did not. Enforcement of the limit was not uniform, and this struck many as unfair for a couple of reasons. Some operating units worked extremely hard to stick within the page limit, cutting and reducing longer documents to fit within the thirty pages, while others simply sent in much longer documents. Survey comments revealed that some operating units that did comply with the limit felt penalized in comparison with missions that

submitted extra information. Technical review committees were given latitude to go back to operating units to request additional information or ask clarifying questions in an effort to “level the playing field” and compensate for the uneven quality and length of the R4 documents. However, because the committees made such requests very inconsistently and unevenly, and sometimes asked for information too late for responses to be received in time to be considered in the scoring, this added to the perceptions of unfairness of the process. There seemed to be widespread consensus that whatever page limitation is imposed, it must be uniformly “enforced” in order to be fair. A number of operating units seem to feel very strongly about this point.

A second issue is that operating unit programs vary widely in terms of numbers of different objectives (including strategic, strategic support, and special objectives), and budget program levels. Imposing the same page length for all missions did not seem reasonable to many of those consulted. In order to remedy this, some respondents advocated using a page limit per SO requirement, so that the allowable length of the R4 would vary in relation to the number of different SOs it has.

There was also confusion over whether the page limit included the data tables or not. Because the required format for the data tables takes up so much space (at most data for two indicators can fit onto one page), there was very little room left for narrative discussion. One person preferred the former style of performance data table used in the API, in which one page could capture targets and actual data for a dozen indicators. A couple of people suggested making the resource request section of the R4 a separate document or separate annex, since it requires so many tables.

(Note: another problem with the R4 performance data table that was mentioned was the difficulty people had in using the format in the software provided. This problem should disappear once the NMS is fully operational. In the meantime, for next year's R4 it is hoped that it would be easier for operating units to use the same format they managed to come up with for this year, if necessary.)

Another issue in the eyes of many respondents was whether it is possible to adequately portray the necessary information within a thirty page limit. A number of operating units commented in the survey that it was difficult to “tell their story” or capture the “true essence” of their progress in the limited space allowed. As mentioned above under adequacy of Agency guidance, the fact that technical review committees asked operating units for more information suggests that the performance information presented in the R4 documents was not sufficient to enable them to do the precise scoring of performance that was needed for the ranking exercises. (This problem would obviously be compounded if there were multiple SOs to be covered in one thirty page document.)

A related issue was raised in comments that stressed that in order to adequately review the performance of an operating unit's program, a technical reviewer needs to have a good background understanding of the program and country context. In this view, the R4 alone

cannot provide sufficient contextual information or detail on performance. This view that reviewers must have background understanding of the program is further supported by the fact that the Agency-wide guidance cable explicitly instructed units **not** to re-explain or re-justify their programs by providing a lot of background information on their strategies in the overview section of the R4 document. The guidance advised units that if no significant changes had occurred in the broader development context, they should briefly note the absence of such changes in this section, and move on to Part II, which describes progress toward objectives. These factors suggest that background knowledge of the context and program does seem to be beneficial and even essential for the review process, given limits on how much and what type of information can be presented in the R4 document. (Being familiar with a program may have the drawback, however, that it can affect the objectivity of a review. A number of operating units complained that review comments were very subjective and inconsistent.) However, some staff commented that it is not feasible to expect reviewers to have in-depth knowledge of all programs, when there are so many different operating unit programs in Africa Bureau (over thirty), and when staff turn over or change positions so frequently. Some operating units apparently felt that not all technical review committee members possessed the needed background understanding of their programs to perform the reviews. Other comments explained that the additional information requested by technical reviewers was sometimes of this general background or contextual nature, and was the type of information that the unit had purposely not included in the R4. They felt that this reflected lack of familiarity on the part of reviewers with the program strategy (CSP) and results frameworks.

**Recommendation:** The following recommendations are proposed to address the R4 page limit concerns.

- Performance data tables should be treated as annexes (or a different format for the data tables that takes up much less space should be permitted);
- a specified range of number of pages should be permitted for the results review narrative for each objective (SO, SSO, or SPO), so that the total number of allowable pages for the R4 is proportional to the number of objectives that an operating unit has, with a suggestion that the range of 3-5 or 3-6 pages per SO for the results review narrative should be the range used; and
- the page limits should be uniformly enforced so that all operating units are treated consistently during the review process.

**Recommendation:** Additional recommendations are proposed to partially address the problem of lack of knowledge of program background or context on the part of technical performance reviewers.

- Copies of the results frameworks graphics for each strategic objective should be included in an annex to the R4 document, to help reviewers understand the program context and for ease of reference.
- Technical reviewers' and desk officers' regular and direct exposure to operating unit

programs should be increased (through more TDYs, participation on virtual teams, etc...).

It was also suggested that an annex containing the full set of performance indicators and data be included in the R4, to assist reviewers who are not as familiar with the program being reviewed.

## **2. Africa Bureau Supplementary Information on the R4**

The Africa Bureau sent out some supplementary information on the R4 preparation and review process, after the Agency guidance cable went out. The first document, titled “Notes on the R4 Guidance” provided information on the R4 process in the Africa Bureau, and took the form of responses to questions raised by operating units. The second document sent to operating units was the draft final report of the AFR Inter-Bureau Working Group, entitled “The AFR FY 1999 R4/PBBS Process: A Proposal and Recommendations.” It contained a detailed description of the review and scoring process within Africa Bureau, including criteria to be used in the reviews. Operating unit survey responses showed that the great majority of respondents felt that the Africa Bureau supplementary information on the R4 was moderately clear or better, and somewhat adequate or better. In addition, a majority of those operating unit respondents who had an opinion felt it was timely.

It was not possible to determine from the survey data which of the two supplementary information documents respondents may have had in mind when answering questions about the usefulness and timeliness of the supplementary information as described above. However, additional free response comments from the survey and feedback from the interviews and focus groups showed that people generally viewed the “Notes” document very positively. It was praised for providing responses to actual questions from operating units, as well as for clarifying some issues. A number of comments from operating units related that staff found the detailed Inter-Bureau report on the mechanics of the process to be too long, complicated, and late to be optimally useful to them in preparing the R4. Some people never even had time to read it, and others felt that the criteria for reviewing and ranking were not clear. (As is discussed later in the section on the effectiveness of the Africa Bureau R4 review process, a majority of respondents to the operating unit survey related that they only “somewhat understood” or worse how all the elements of the review process were conducted. Approximately half of the respondents indicated a level of understanding of the process at below “somewhat” understanding it.) Another relevant point that is mentioned above in the discussion of the Agency-wide R4 guidance is that some people complained that instructions kept changing and were conflicting (and caused them to have to keep revising their R4). This perception may reflect the fact that three different documents in total were sent out (guidance cable, “Notes”, and Inter-Bureau Working Group document). Finally, some staff felt that because information was tailored for sustainable development country programs, it was less useful to programs in other categories.

## **3. Success of the R4 in Meeting its Objectives**

As described above, assessment designers assumed that there are a number of different purposes for the R4, and multiple views as to what those purposes are. In light of this, the assessment tried to determine whether staff felt that they clearly understood the objectives of the R4, how they would articulate those objectives, and whether or not specific objectives were met.

**a. Clear understanding of the objectives for the R4**

The questionnaire asked operating units whether the primary purposes or ultimate objectives of preparing the R4 and reviewing it internally within the operating unit were clear. The great majority of those who responded answered “yes” to this question. When asked to describe the purposes for the R4 preparation and internal review process within the operating unit, many mentioned assessing progress toward objectives or results. (As will be discussed below, the majority of operating unit respondents felt that it was successful for this effort.) Other purposes of the R4 preparation and review process within the operating unit that were mentioned were: reviewing performance as a basis for allocating financial resources; matching resources with program needs; verifying or revising the management contract; reviewing the need to revise indicators and/or results; using information on results to take management actions such as making implementation adjustments and reallocating resources, or using it as a management tool; reporting results to USAID/Washington, and larger audiences such as Congress and the public; securing resources needed for achieving approved CSP objectives; providing information to Washington to help them allocate resources; and ensuring that partners and stakeholders have a clear understanding of results and how they are being reported. A few people interpreted the question on the purposes for internally reviewing the R4 more literally and responded that it was to comply with Washington directives and guidance, or for things like quality control, verifying data, making the R4 document internally consistent, increasing staff understanding of programs across the operating unit, and ensuring staff consensus and ownership of the document.

A majority of operating unit survey respondents also indicated that the primary purposes of the USAID/Washington review of the R4s were clear, although a significant minority felt that the purposes were not clear. This suggests that operating unit staff were slightly less clear about the purposes of the Washington review of R4s than they were about the objectives for the internal preparation and review process within the operating unit. This may be linked to lack of complete understanding of the details of the Africa Bureau review process by operating unit, primarily field-based, staff.

Respondents to the operating unit survey described the purposes of the USAID/Washington review of R4s as to help Washington accomplish the following: understand and review the performance of operating unit programs and their progress toward results; use performance and other criteria as a basis for deciding future budget allocations; rank operating units based on performance, with the implication that this will inform inter-program resource allocation; compare program performance to the management contract and

validate or revise it accordingly; learn about the operating environments and implementation issues affecting programs; provide feedback to operating units on their performance, including recommendations on monitoring and program management; promote dialogue and understanding on programs between USAID/Washington and operating units; identify areas where programs need technical assistance; and provide information for reporting to Congress on U.S. development assistance.

**b. Success of the R4 in terms of the objective of assessing progress toward results**

One of the primary objectives of the R4 preparation and review process is to enable USAID to assess progress towards results. Feedback from the interviews and focus groups of Washington-based staff revealed that there was a general sense that the R4 was successful for the general purpose of assessing program impact. The operating unit questionnaire asked staff about different specific dimensions of assessing progress towards results. The great majority of operating unit respondents indicated that they thought that preparation and internal review of the R4 was moderately successful or even better as a means for the operating unit to assess and evaluate its own progress toward results. A number of positive comments were received through the operating unit survey about the preparation and review of the R4 by the operating unit to assess its program performance. The process helped staff get informed about program status, increased information sharing across strategic objective teams, reinforced teamwork, and helped partners understand the emphasis of reporting on results as opposed to outputs. Some operating unit staff liked the fact that the process forced them to assess their progress toward results. It involved using a more systematic and intense data collection, analysis and presentation approach, and helped units identify where performance tracking systems needed improvement. Several staff mentioned that the internal review of performance was very helpful in their development of a new strategic plan, since lessons learned from past performance were used to redefine strategic objectives for the new plan.

There were also some negative comments on the success of using the R4 for the internal performance review process within the operating unit. Some staff felt that they spent too much effort on collecting data, and not enough in analyzing it. Others felt that the emphasis in the preparation and internal review of the R4 was too much on the intended Washington audience and not enough on doing a comprehensive review of performance at the operating unit level. Some mentioned that preparing the R4 took valuable time and effort away from more in-depth internal reviews, or that the R4 drew largely from the semi-annual results review. Semi-annual activity reviews were praised by a couple of staff as a more useful management tool, one that could have achieved the same objective as the R4 but more efficiently. Some operating unit staff reported that their unit did not have an internal review as part of the R4. One person related that their unit was too busy at R4 preparation time, and found it more useful afterwards for working with their partners. Some staff felt that strategic objective teams already knew the status of their programs, although one person added that even though they already knew whether they were on track vis a vis meeting indicator targets,

the R4 presented a useful chance to step back and review the big picture of performance. In some cases, good data were lacking, or were not produced by the monitoring systems in a timely fashion.

A number of operating units had to prepare a new Country Strategic Plan at virtually the same time they were preparing the R4, and some problems were associated with this. Staff were hard pressed to do both at the same time, and it caused a very heavy work load for operating unit staff and desk officers. In at least some cases, country desk officers were not able to fully participate in R4 discussions because of CSP competing needs. One mission requested special permission to prepare a briefer R4, but did not get a clear answer from Washington. In some cases either the CSP or the R4 received short shrift, within the operating unit and/or in the Washington review process, (comments suggested that CSP preparation and review seemed to take precedence over the R4 in these cases). There was also some conceptual confusion as to how to deal with both the review of past performance based on the ending strategy, and the review of the new strategy at the same time. Comments suggested that this confusion affected both operating units, who were uncertain of exactly what to report on in the R4, and Washington staff, who kept referring to the new strategy in their review of past performance as presented in the R4. While the Africa Bureau did have an exceptionally high number of new CSPs to review during the FY 1999 R4 preparation and review season and this should not be the case again for at least a couple of years, the assessment feedback does suggest that the issue needs to be addressed in the design of the R4 review process.

The operating unit survey also asked about the related issue of the R4 document's success in providing information to USAID/Washington on the impact of the operating unit's program. The majority of operating unit survey respondents indicated that it was moderately successful or better at this. Assessment findings also give the overall impression that Washington-based staff thought the R4 did work well in the area of assessing program impact. A benefit mentioned by several people was that it gave Africa Bureau staff an opportunity to review and become familiar with a great number of different operating unit programs. However, there were also quite a few comments to the effect that the R4 document was not completely successful in providing information on impact for the AID/Washington reviews.

Constraints to success in providing impact information to Washington that were cited in the operating unit survey comments included lack of clarity within the operating unit on what information should be included in the R4 report, and selective reporting on only some indicators as called for in the R4 guidance as well as the R4 page limitation that many staff felt prevented units from effectively communicating the full story of impact. A number of staff thought that dialogue between USAID/Washington reviewers and mission representatives was extremely helpful and even essential during the review process, since the limited information contained in the R4 document could not present the full picture of impact. Also, as discussed earlier in the section on the Agency-wide R4 guidance, the fact that technical review committees asked operating units for so much additional information suggests that the

R4 document may not have provided all the impact information needed for the review process. Some people thought the document provided good impact information, but felt that they either did not understand or did not agree with how USAID/Washington staff interpreted or used the impact information, and that review decisions were inconsistent. Some Washington staff thought that the highly varying quality of the R4 documents was a determining factor in the success of the R4 in providing impact information to USAID/Washington. They commented that some operating units wrote excellent documents, and others did not. Another issue raised was that many of the results frameworks for the strategic objectives in operating unit programs do not contain strong causal linkages between lower and higher order results. The argument was that since the causal links represent the development hypotheses about how impacts are achieved, if these are not persuasively causal, it is not possible to effectively judge whether the strategy has been successful in achieving the strategic objective or not.

A number of people were concerned that there is an overemphasis in the R4 document on relying on a few, usually quantitative, performance indicators to try to report impact, to the exclusion of more qualitative and anecdotal types of information that could be best captured in narrative form. The concern that more qualitative and narrative information are needed to supplement the few quantitative indicators in order to give the full story of impact was voiced by both operating unit and USAID/Washington staff. People felt that the R4 page limit contributed to the lack of opportunity to include more qualitative information in the narratives. Several mission staff commented that expectations for quantifiable indicators in the democracy and governance sector particularly are unrealistic and impractical, and advocated for greater acceptance of qualitative indicators in that sector.

Several staff raised the issue of the complexity of the development process and questioned whether it can be adequately captured in performance reporting, with the type of information being reported in the time frame allowed and formats required for the R4. They argued that complex development changes do not easily lend themselves to representation through a few indicators. Progress occurs very slowly over time, particularly at the higher impact levels of the SO, and trying to capture annual change is often difficult or impossible. A few people suggested that performance reporting should occur only every other year. Some saw a conflict between the longer term goals of performance analysis in capturing development progress, and shorter term reporting needs mandated by the GPRA and Congressional oversight. In a similar vein, concerns were raised about the difficulty of reporting on newer strategic objectives. The Africa Bureau R4 review process required that operating units report on all SOs that had been in place for at least eighteen months, but some people felt that eighteen months might be too soon to show progress for new SOs, given that many months are often needed to get activity implementation mechanisms in place (see later discussion on reporting on newer SOs in the section on how well the review process worked).

A related concern was expressed that the R4 may not be providing the types and richness of information needed to meet the Agency's external reporting requirements to

Congress and other external audiences. In particular, some people felt that it is not providing sufficient information for the Annual Performance Report. It was suggested that the R4 might not be the appropriate vehicle for providing the range and depth of information needed for reporting to Congress.

**Recommendation:** The following recommendation is proposed to address the issue of improving the success of the Africa Bureau's R4s in providing impact information on programs to USAID/Washington.

- Good examples of R4 narrative discussions should be circulated to all Africa Bureau operating units at the same time the R4 guidance is issued, along with supplementary materials discussing how to successfully present qualitative information on impact.

**Recommendation:** The following recommendation is proposed to address the difficulties caused by simultaneous preparation and review of country strategic plans and R4s.

- The R4 guidance should include clear instructions for how to deal with both content and process issues arising from submitting a CSP and an R4 at the same time.

**c. Success of the R4 in terms of the objective of using performance information to influence budget allocation decisions**

Using performance information to help make budget allocation decisions was cited as a primary purpose for the R4 process by a great number of operating unit and Washington staff. However, it was not considered to be very successful for that purpose by the majority of those consulted. Over half of those who responded to the question in the operating unit survey thought that the R4 preparation and Africa Bureau review process was less than moderately successful in making sure that information on progress toward results was used as a factor in making decisions regarding budget allocation for their operating unit's program. This issue seemed to be a big concern in the eyes of many.

While a couple of positive comments on the use of performance information to make budget decisions in the Africa Bureau R4 process were made, most comments received were negative. A number of survey comments revealed that operating units saw little evidence that performance influenced budget decisions at all. One person commented that their operating unit OYB increased as a result of good performance impact achieved. In contrast, a number of cases where cuts were imposed despite high performance, and where low performing SOs did not receive budget cuts, were cited. The Kenya and Ethiopia programs were specifically mentioned by numerous people as examples of this. Political, foreign policy, and other factors besides performance clearly predominated in these decisions. Many criticized the fact that performance only counted for 35% of the total score, and there were suggestions to increase the performance share. Many staff did not understand how the 65% non-

performance part of the scoring was calculated. There seemed to be a sense that linking budget decisions to performance was important as an incentive to improve performance. Many staff felt that the real budgeting decisions are still made behind the scenes, without operating unit input. People wondered whether all the time and effort going into the R4 preparation and review process was worth it, given their perceptions that it had so little effect on resource allocation.

People recognized in their comments the great obstacle to performance-based budget allocation posed by fact that Congress is wedded to the older sectoral earmark and projects model, which they felt conflicts with the basic intent of the GPRA legislation. There were a number of calls for greater transparency and explicit description of how results are related to resource allocation in the Africa Bureau review process. Some wanted greater operating unit involvement in these decisions. It was suggested that if the decision-making were made more transparent, the R4 could be better geared to help make those decisions. The issue was raised that most SO performance scores ended up being clustered around a narrow range, while the non-performance part of the total scores displayed a much wider range of values. If valid, this could have lessened the relative effect of performance on total scores to an even greater extent. One person thought that mandatory common indicators across programs would be necessary for real performance-based budgeting to work, so that standardized data would be collected and programs could be more easily compared with each other. This would require programs to be designed from the beginning to accommodate the common indicators, which cannot easily be imposed after a program is already running.

Some staff suggested that performance ended up confirming tentative budget decisions, since most programs received close to their requests. One person speculated that linking budgets to performance may work best in an environment where resources are less constrained. A few staff recognized that the budget allocation process was not completed at the time of the assessment, and felt that it was premature to try to judge this issue at that time. One respondent felt that the budget allocation process could never be made completely standardized and objective, and so instead of trying to develop a complicated budget allocation model, senior managers should be given some flexibility to make the decisions.

As alluded to above, a limitation of the assessment feedback on this subject is that the assessment data were gathered before the final resource allocation decisions were completed in Africa Bureau. Initially the Bureau's budget level was set at \$718 million, close to a general equilibrium level where most programs could receive what they were requesting. At this budget level the effect of performance on resource allocation decisions was modest. However, as the Bureau developed budget scenarios at the \$702 and \$670 million levels, performance had a more significant effect on allocation decisions. The Africa Bureau has received praise from OMB for the way that it incorporated performance data in decisions. Most operating units, particularly in the field, may not be aware of this. Informing them of how the final budget scenarios played out may be important as new guidance for the coming cycle is considered and developed.

**Recommendations:** The following recommendations are proposed to address the widespread concerns about the use of performance information to influence budget decisions.

- The Agency needs to clarify what the real relationship between performance and budget allocation decisions is. It needs to determine what the appropriate percentage for performance should be in overall budget ranking scores.
- The 65% “black box” or non-performance part of the scoring should be made more explicit, transparent, and readily explainable.

#### **d. Managing for Results and the R4**

The Automated Directives System does not specifically require that operating units present evidence that they are managing for results (MFR) in their R4 documents, nor does it specifically mention that MFR will be an explicit factor in the review of performance for the unit. The FY 1999 R4 Agency-wide guidance cable, however, did instruct operating units to include a description of management actions the unit is taking or will take in response to performance assessments for objectives in the R4 document (they were supposed to include this discussion in the “Part II: Progress toward Objectives” section of the R4 document). The Africa Bureau supplementary Inter-Bureau Working Group report also specifically mandated that managing for results would be a factor in the SO performance reviews. Managing for results was defined for this purpose as “evidence that performance information is utilized to verify program content and direction and to revise program activities and/or results frameworks as necessary.” Later, technical review committees were instructed to allocate 25% within the total 35% performance score for an objective to considerations of MFR.

The majority of survey respondents reported that their operating units have used information from the R4 preparation and internal review process to make program decisions or manage for results, and many provided specific examples of how they had done so. Some stated that their units were using performance information to make decisions anyway, as part of an ongoing monitoring and evaluation system or through semi-annual operating unit results reviews, and not specifically as a result of having to prepare the R4.

An interesting issue that arose in assessment feedback is that there is not a clear consensus and shared understanding of what managing for results means and how to identify it among staff within the Africa Bureau. Respondents identified a wide variety of types of actions as evidence of MFR. Some technical review committees apparently felt that there was sometimes not enough information contained in the R4s to judge MFR issues. There seemed to be a feeling that it would be helpful for both operating units and technical reviewers to have more explicit guidance on what managing for results is, how to report on it, and how to score it, so that the review process could handle MFR issues more consistently. Examples of some of the wide ranging types of things that people identified as evidence of managing for results included: having SO team structures in place; improved team interaction; having data collection and reporting systems in place or improving them;

discussing with partners the need for improved data collection and reporting, or meeting with partners to prioritize interventions; identifying or revising indicators; revising intermediate results, strategic objectives, and/or targets in response to performance data; adding or cutting elements of an SO strategy; designing new programs or elements of programs in response to past performance information; refocusing or adjusting activities; revising funding levels for specific grantees; and allocating budget resources, or adjusting resource requests for an SO. It is interesting that some of these fall more into the area of operations and how the Agency and its partners go about their daily functions, rather than actually involving making decisions about program directions in response to performance information.

An important point raised is that earmarking is fundamentally contradictory to effective managing for results. The funding restrictions imposed by Congress interfere with the ability to truly manage in a performance-based manner. Another interesting conceptual issue is that typically managing for results is identified through evidence that adjustments have been made to programs in response to identified problems. The question was raised as to how an operating unit can demonstrate (and how a technical reviewer can identify) that it is managing for results when the program is going well, and no program changes are required. An additional point from the assessment feedback was that managing for results is based on the assumption that the operating unit will be honest in reporting results, and not hide problems. However, some operating unit staff felt that honesty in reporting led to programs being penalized in the reviews, while good “public relations” reporting in the R4 was rewarded. In addition, the concern was raised that being honest about problems could hurt the Agency in the aggregate, given “hostile” attitudes towards Agency programs among stakeholders in Congress.

**Recommendation:** The following recommendation is proposed to help clarify issues affecting managing for results and the R4.

- Clear consensus on what is meant by “managing for results” should be developed. Good examples of demonstrating managing for results in the R4 should be circulated to all Africa Bureau operating units at the time that R4 guidance is issued.

**e. The feasibility of meeting all of the purposes for the R4**

As discussed earlier, there are a number of different purposes or objectives for the R4, as well as a number of different views or interpretations of what those objectives are among Africa Bureau staff. To some extent, objectives for the R4 preparation and review process can be analyzed from an Agency, Bureau, or operating unit perspective, although there is considerable overlap and commonality among these perspectives. A finding of the R4 assessment is that there seems to be a need for more discussion and clarity around the multiple purposes of the R4 in the Agency as a whole, as well as within Africa Bureau. The question is raised as to whether it seems feasible for the R4 to fulfill all of the purposes or expectations for it. Is the Agency trying to accomplish too many things with one document,

given the limitations on how much and what kind of information can be presented in the R4? Are the resource-intensive performance data collection, analysis and reporting systems that are required sustainable in an operating environment that is characterized by continually reduced Agency staff levels? Can both annual and longer-term, as well as internal and external reporting requirements be met with one document, particularly given the inherent difficulty of trying to capture the effects of assistance on long-term, complex development change processes? Can the R4 be successfully used to allocate Agency budget resources, given the reality of the funding constraints imposed by Congress?

A number of staff commented that variations in R4 quality were dramatic, and that this affected the utility of the document for various purposes. Some Washington staff who were involved in the review process felt that the capability of the person who actually wrote the R4 document had an enormous effect on the quality and usefulness of the document for the review process. Some knew how to report well, and others apparently did not. The related issue of honesty in reporting was also felt to affect the usefulness of the R4 documents. While some operating units reported openly about negative progress, others allegedly chose only to write about the best elements of their programs. Some operating unit staff apparently thought that honest reporting led to programs being penalized in the reviews, while those who used the R4 more as a “public relations” reporting tool were unfairly rewarded. In some cases where technical reviewers knew programs well they apparently questioned data they thought were suspiciously “rosy,” but this did not always happen or there was sometimes insufficient time before program review day to do anything about it. The issue was also raised of whether transparency in R4 reporting could cause problems with host country governments, in the event that negative information about the government was included in the document. This prompted a question about how widely the R4 document can be shared with different audiences. In summary, there seemed to be widespread sentiment that the quality and transparency of the R4 is critical to the usefulness of the document for its multiple purposes.

**Recommendation:** The following recommendation is proposed to address the issue of the multiple purposes of the R4.

- The Agency needs to clarify and be explicit about the purposes and objectives for the R4. It needs to examine whether the R4 can meet all of its current objectives.

## **B. HOW WELL DID THE AFRICA BUREAU R4 PROCESS WORK?**

Responding to all aspects of how the process worked captured the interest of the respondents and elicited a large number of comments in the assessment. Everyone has ideas on how to make this work better or at least where to attack the problems they do perceive. Almost no one was neutral about discussing the process. Most saw the review process as valuable as a learning opportunity for Washington and still had serious reservations on its

utility as a budget setting exercise. Opinions on the performance scoring seem to have been colored by whether they thought Bureau decisions were ultimately informed by the scoring or whether they perceived other variables as being more important. While feedback did not attribute subterfuge to anyone, many clearly want the issue of game rules and playing field dimensions to be dealt with forthrightly.

The following aspects of the Africa Bureau R4 review process were praised: application of a consistent set of criteria; operation of a transparent and rigorous system for scoring; the participatory nature of the process; increased information-sharing and knowledge as a result; the efficiency of program days; dedication and hard work by staff and the responsiveness of key DP

staff to queries. In addition, the competence of the technical review committees and the quality of dialogue between them and the operating units was appreciated by many operating units.

Numerous aspects having to do with the effectiveness and efficiency of the Africa Bureau R4 review process were singled out for critical review and specific improvement. These will be presented and discussed in the following section. Recommendations to address these issues are presented as a complete set at the end of the section.

## **1. Budget Categories in the Performance Scoring Process**

The budget category-based system of scoring performance did not work well for assessing strategic objectives that were funded from more than one budget category. The majority of respondents in the operating unit survey rated this part of the review process as below moderately successful. Breaking population/health/nutrition SOs apart into separate components of HIV/AIDS, child survival and population as well as splitting apart SOs involving combined economic growth with democracy/governance or environment/natural resources management funds were cited as particularly problematic. Almost no one seems to think that the way that this challenge was treated worked well. There were a number of specific criticisms.

- Operating units responded that synergies between sectors were not analyzed well, and the budget category-based analysis system was less successful overall. Comments noted that synergies are seen as obstacles to technical rating concerns, that aspects of mutually supportive strategies such as health-related activities were lost, that operating units received little information on how synergies were considered, and that linkages were not explicitly looked at and some felt were even deliberately ignored. Artificially breaking strategic objectives apart into sub-components led to forfeiting a technical assessment of the SO as an integrated whole and made it difficult to capture synergies at all.
- When more than one funding source is being used by an SO, many thought that the

predominant emphasis of the SO should be scored by technical reviewers who know that sector the best. Conducting separate reviews for each sub-component required more time and effort. This led to the use of multiple committees working on the same thing, and more requests for extra information from operating units. It was suggested that some aspects of an SO that are outside its predominant category (e.g. IRs which support the SO and come from different sectors) could be scored by asking for specific resource person input from another sector team as needed. This could be defined and requested and could avoid lots of duplicate effort. Setting up a working group ahead of the review process to decide who should take the lead on which SO and where additional expertise might be needed as a supplement was suggested as a potential remedy for the confusion and double effort.

- Heavy inclusion of other bureau funds (such as BHR) in humanitarian assistance programs was an extra challenge for the reviewing teams. Additional information was needed not from the operating unit but from central bureaus in order to complete scoring. This lack of available information reduced the numbers of SOs in the humanitarian assistance area that could be successfully scored.
- The process was inherently unfair to operating units whose programs were being judged on the basis of categories and criteria established after the fact. Since they were not required to delineate separate sub-sectors in strategic planning and performance measurement systems at the planning and achieving stages for multi-funded SOs, how could performance be legitimately judged on the basis of separate categories later?

A related issue is that the sectoral focus of the review process was seen as beneficial but working against looking at the overall program impact picture. How to focus on analyzing and comparing SOs in a fair and comprehensive way across the continent versus how to look at each country program's progress as more than the sum of its parts was seen as a particular challenge. If all the elements are looked at individually, is there a way to see whether the total impact equation merits the resources being used?

## **2. Complexity of the Review Process**

The review process is seen as overly complex. People described their views of this complexity with numerous examples. Asked how well they thought they understood how all the elements of the review process worked, approximately two-thirds of operating units respondents to the survey said that they only “somewhat” understood it or worse. Many aspects were judged to be unclear or overly complicated. There were some specific concerns.

- Field operating units did not know what the technical review teams were being asked to do and what the budget level setting process was about.

- Milestones in the review process would help everyone understand what point the reviewing had reached and what would be happening next.
- The numbers of SOs each team had to work on greatly affected both efficiency and ability to communicate with the field as needed. Some teams had a few SOs, some had dozens. Sufficient time to do reviews is a problem when there are large numbers of SOs to handle.

**a. Performance scoring process and criteria**

The performance scoring process itself generated large numbers of comments. Most found the process transparent but overly intensive. The degree of complexity in the instructions and the processes used led some to view the process as exemplary in its effort to be objective but nevertheless cumbersome.

While everyone seems to want consistent and clear scoring criteria, many challenged whether a ten point scale was needed to accomplish the ranking tasks at hand. The scale was seen as time consuming to use and led to the pursuit of “false precision” as teams struggled to refine score to tenths of a point. Could a simpler scoring system, such as one based on fewer categories like “high,” “medium,” and “low,” serve the Bureau's needs as well and also take less time? Operating units commented that a simplified scoring process would allow more time for technical reviewers to suggest ways to improve impact. The following key points were raised.

- In the ten point scale, what is the rationale for setting six as meeting targets? Does this set up an incentive for operating units to try to always exceed targets? Will this affect the ambitiousness of target levels over time?
- The mean score in a sector was assigned as the “neutral” score in cases where there was not enough information to judge performance. While the intent of this procedure was to neither penalize nor reward new strategic objectives, it was pointed out that it might provide an incentive to operating units to withhold negative performance information in order to try to avoid a low score and instead receive a higher (neutral) mean score.

Additional problems with the scoring process that were cited included: excessive amounts of time spent discussing individual scores at times; variations in scoring between G Bureau and Africa Bureau team members leading to contention in some sectors that is not fully resolved; presence of field unit advocates on teams who scored higher because of their familiarity with the programs under scrutiny; apparent subjectivity in some of the scoring; M Bureau preference for quantitative scoring; false impression that more precise scoring would help budget decisions later; lack of transparency in the process used by the D/G team; if wrong indicators are being used, then emphasis on scoring will not help the program improve;

weighing indicators equally in the scoring when their relative weight in the results framework might be more instructive; subjectivity is not removed by using weights or rankings; and viewing the process as a science instead of an art.

An interesting issue concerned how well people thought the scoring process worked for comparing performance across countries. While some staff thought the review process worked well for this and that it was very good for the Agency to begin to review resources globally, a number of operating unit staff were concerned about it. They thought it was unfair to try to compare programs in different countries where the environment, host governments, and contexts vary so greatly, and that each program should be evaluated on its own, according to its strengths and weaknesses. The point was also raised that a logical implication of comparing performance across countries is the need to move toward more standardized reporting using common indicators, which will mean more restrictive program parameters and designing programs a priori to accommodate the standardized indicators. (It does not work well to try to impose common indicators after a program is already operating on the ground.) There is some concern that this standardized programming and monitoring approach is counter to the spirit of reengineering and managing for results, and represents a more centralized, vertical control approach to program management.

**To sum up, the Washington respondents seem to be saying: make the scoring process simpler so we can spend time talking about how to improve performance. Field units seem to be saying: make it simpler so we can understand how we are being evaluated.**

### **3. Feedback and Communication between Washington and Operating Units**

Feedback from USAID/Washington and dialogue between Washington and operating units were viewed as needing improvement. There were some communications successes cited, notably the way the environment review team dealt with the field. Even those units which did not like the process thought they learned something from it just the same. Nearly half of operating unit survey respondents thought that the R4 review process was less than moderately successful as a mechanism for USAID/Washington to provide feedback to the operating unit on performance. A number of critical concerns were voiced.

- Technical performance scores and explanations of those scores, final composite scores, and reporting cables were sometimes sent to operating units late, or not at all. Operating units voiced frustration with transparency issues and with the lack of explanation for the scores they did receive. Many expected some analysis with which to advance the dialogue on performance and did not get it. Of those responding, many expressed their dismay that feedback was still pending at the time the survey was completed (late August- early September 1997).
- Some technical review scores were not ready or were not shared with operating units

far enough in advance of program day.

- Technical review committees asked operating units for more detailed information on performance than the R4 guidance required or even allowed (given page limitations). This was both advantageous and burdensome to operating units. While it enhanced understanding of programs and dialogue with operating units, committees did this inconsistently. This was cited as unfair to those units that did not get the opportunity to present additional information. It was suggested that some form of teleconferencing between Washington and field units and/or vetting of R4 information with virtual team members could lead to a more equitable interpretation of what was reviewed.
- Late timing of questions, multiplicity of requests for more information, and requesting inappropriate types of information were raised as problems by operating units. Some thought that requests reflected lack of familiarity on the part of reviewers with program strategy (CSP) and results frameworks. Time frames for responding to requests for additional information were insufficient in some cases and the complexity of the additional material sought proved annoying to some units queried. Some would have preferred to be rated on what was sent in only.
- Some operating unit staff would like more opportunity for dialogue with USAID/Washington about scores during the review process and to influence scores before they are finalized.
- A few people thought that there is no appeal procedure built into the review process for operating units to appeal scores.
- Assessment findings imply that even though information may have been sent to operating units, it was not always shared widely with staff within the units. It seems that Africa Bureau supplementary information, as well as feedback and scoring information from the technical reviews, was not always distributed to all staff within an operating unit. Scores and ranking information that were sent to geographic offices in the Africa Bureau were apparently not always distributed widely to Washington-based staff either. This suggests that operating units may need to work on improving internal communication and information-sharing channels.

Remarks also dealt with dimensions of what people called “**leveling the playing field.**” This issue had to do with the fact that technical reviewers often requested additional information from operating units to supplement what was presented in the R4 document. There are two schools of thought and both are represented in the reviewing groups and in the operating units. One school says: it is good to get more information when something is not clear or seems inadequate and everyone does a better job as a result. The other school says: make it clear from the beginning and evaluate everyone the same way from what they send in. These points of view are probably not reconcilable and do affect a person's perception of the validity of the performance rating process no matter what. Several critical issues are

related to this divergence of view: whether one thinks page limits should be strictly adhered to, whether a field unit thinks lengthy queries from review committees are attempts to be fair and helpful or not, whether the presence of technical expertise familiar with the country on the committee is a good thing or not, etc... In seeking ways to streamline the reviewing process both aspects - fairness and collegiality - must be given serious attention.

#### **4. Technical Review Committees**

The fact that some technical review committees went beyond the mandate of the R4 review and also analyzed and criticized program strategies was a sore point. Sometimes this was seen as a lack of familiarity with existing programs and decisions; sometimes this was seen as exceeding the brief for the review process.

Concerns were also raised about the way that review committees operated. Did they read all the documents? Was every field unit asked for more information or only certain ones? Did committee membership remain constant throughout the process? Were operations transparent? Was enough time spent writing up the team's results? What kinds of processes were used to reach agreement?

As discussed earlier, concern was expressed about level of country knowledge and lack of knowledge of key field unit documents such as the CSPs on the part of technical reviewing groups. Sufficient in-depth knowledge of programs was seen as beneficial and even essential for reviewers to have, especially given the limits of how much and what type information can be presented in the R4 document. There was discussion about how to address this including finding ways to legitimately involve CDOs in the process to offset any lack of knowledge on a particular country program, more regular visits to operating units by country desk officers with the explicit intention of discussing performance issues with all the SO teams, more systematic deployment of key technical staff to a range of operating units for short-term assistance, and insistence on the digestion of background information on each program by the review committee members. While increased knowledge was seen as good for all concerned in the process, it was also recognized that it would affect the objectivity of the reviewers in some way. This effect could be both positive and negative in its outcome as far as performance review was concerned.

Flexibility in the way that the review process functioned in the individual technical review committees was seen as a positive aspect but not without its down side. Letting each group develop its own modus operandi within the task limitations probably worked well in the long run but may have been more time consuming. Suggestions for improving this process include: team-building coaching for the groups before they start work; norm-setting exercises to govern how committee members will each work together and make decisions; grounding key technical staff so that a critical mass is maintained for the duration of the group's work; finding ways to work together virtually to cut down meeting time; consulting staff who track certain countries before the team starts deliberations to fill in any knowledge gaps that can be met without going to the field for information; and using the CDO as a

resource person when appropriate.

## 5. Newer Strategic Objectives

How to deal with newer SOs in the review process was raised by Washington and the operating units in the field. It is noted that there are currently a large number of new approved strategies so this issue will become even more important next year. While the guidance did specify that SOs that were eighteen months or older should be reported on and reviewed, more detailed and practical parameters for when to look for performance information from these SOs seems to be called for. **People want to know when is the best time to start expecting performance information on these SOs.** This issue is related to overall strategy maturity. It is often not just the SO that is new but perhaps the whole approach and combination of SOs.

A corollary concern was the lack of information on funding levels for existing SOs. It was judged to be hard to review impact and progress if one was not sure about how much should be expected from the resources in play.

## 6. Program Review Days

Operating unit representation is seen as critical to “healthy” program review by many staff. The majority of operating unit survey respondents thought that the unit had adequate opportunity to participate in the review process. Approximately twenty-three out of thirty-two units were able to participate in a program day. Most of those commenting think that all units should have a chance to be heard directly or none should. There was the suggestion that limited programs should be given priority if all are not represented since less is known about their programs than for the sustainable ones. Transition countries are seen as needing face to face review as well since most of the Bureau's outside requests for information on performance and direction of strategies are about these countries. Limited missions were also viewed as needing full program days. Mission director participation in program day reviews was seen as critical as well. Some units requested the presence of SO team leaders to more fully explore the dimensions of the performance information.

Program days were positively viewed in comments. Suggestions were made that program days would work better if there were more opportunity for discussion and the meeting did not primarily focus on review information feedback. They would like the larger issues of program performance addressed and found that some program days did this well. There was a request for more dialogue rather than just R4 report discussion. The R4 could therefore be the means to the end of a comprehensive look at a country program but not the only focus for it.

## **7. Senior Management Participation in the R4 Review Process**

Participation of senior bureau management in the review process was raised repeatedly. The managers themselves voiced their desire to be able to participate more comprehensively and other staff expressed their wish to have them present. It seems that explicit formats and means must be found to increase their presence in the process. Full time allocation will not work with the demands outside the Bureau. Suggestions included: AA meetings with each field unit in a condensed format; weekly two hour wrap-ups for senior managers that summarize key aspects of review process progress that week; and one hour sum-ups at the end of each program day where key decisions can be discussed and key issues tabled with the senior staff.

## **8. Efficiency of R4 Preparation and Review Process - Was the Level of Effort Appropriate?**

More than half of the respondents in the operating unit survey indicated that they thought that the amount of resources spent on preparing the R4 was too much, while roughly a third felt that the amount was about right. More than a few thought that the time spent answering Washington's additional questions was not reasonable. Some noted that the R4 preparation process provided an opportunity to think seriously about results and also that current staff levels will not allow the continued allocation of this kind of effort without some streamlining. A few yearned for a process which was year-round in its data collection efforts in order to allow for a more expeditious preparation period when the R4 is in its final phase before submission. They question whether an annual process will really work and ask that a biannual time frame be considered. This would include annual reviews within each unit and biannual submission to Washington.

A representative comment would be: "Given what it was used for and the impact on the operating units, sleepless nights and 18 hour days for some staff members, do not justify the outcome or impact". An operating unit complained that staff spent considerable time and effort in reducing the information they had prepared for the R4 to the 30 page limit, only to be asked later by technical review committees to provide the very information that they had cut out of the document with such difficulty. Considerable level of effort was required to respond to multiple and at times competing or repetitive requests for information from technical review committees. Some units mentioned that the great emphasis on preparing and following up on the R4 document for the Washington review process took time away from valuable internal program review within the operating unit.

Assessment findings are that almost all Washington-based staff consulted thought that the level of resources spent on the Africa Bureau R4 review process was too much. Nearly two thirds of operating unit respondents indicated the same. The level of effort used in the review process was described as extremely high and posed a significant burden on staff that is not sustainable. As one operating unit put it: "we submitted R4... on March 8, program day was May 5, ... reporting cable was sent...August 18! Suggest simplify it, shorten it, make it

more transparent...” While many viewed the program days as efficient, well-conducted, and a good use of time, some also noted that the intensity of the review process limited opportunities to interact with the operating unit staff while they were in Washington and that this was an opportunity lost.

Large numbers of comments focused on the human cost factor for both preparation and review. It was noted that this aspect must be taken very seriously given the reality of reduced staff levels in the Bureau. Some people seemed to think that with practice the process might become more efficient but the reviewers noted little evidence beyond wishful thinking to support this view. Now that people understand more about what the outcomes of the review mean, there will in all likelihood be even more attempts to put a fine point on discussion and decisions, probably adding to the time need to reach conclusions.

In terms of the requirements for limited missions it was suggested that a more condensed reporting and review process might be warranted. This could be considered along with ways to reduce the number of units reviewed every year in Washington by going to a biannual schedule with an internal review annually in each operating unit and a report on any changes which affect the management contract at that time. Some noted that although the R4 process is seen as lengthy and time-consuming it is a question as to whether it is more so than the previously used discrete processes: PIRs, API, ABS, and CPSP, not to mention the project design and approval cycle with its panoply of documents. While staff seemed to think that the Africa Bureau R4

performance review process was valuable and merited significant effort if it were to be used for resource allocation decisions, many wondered whether it was worth it.

## **9. Recommendations on How to Improve the Effectiveness and Efficiency of the Africa Bureau R4 Process**

**Recommendations:** The following recommendations are proposed to address the issues raised above concerning the effectiveness and efficiency of the Africa Bureau R4 process.

- Clear, concise supplementary information on the mechanics and details of the R4 review process should be sent to operating units (the simpler it is and the earlier it is sent the better).
- The performance scoring system for SOs should be revised so that:
  - a simpler system with only three or four categories such as (high, medium, low) replaces the ten point scale, or
  - if the ten point scale is maintained, the assigned values for “meeting targets” and “neutral” are revisited.
- Procedures for reviewing performance of SOs funded from more than one budget

category should be revised so that:

- a single sector technical review group takes the lead on each SO, and
  - there is no double rating of any SO.
- The review process should be streamlined. Rating deliberations and requirements should be reduced. Potential ways to do this include switching to a simpler scoring system as described above, and/or utilizing smaller preview groups that prepare preliminary ratings to share with larger technical review groups for final decisions. A task force should be created to examine ways to increase efficiency and streamline the review process. ANE Bureau should be consulted for information on how they streamlined their process this past year. Whatever system is devised, it should stress maximizing efficiency in order to accommodate the reality that Africa Bureau has a greater number of operating units than any other geographic bureau.
  - Consider conducting a full review of the R4 for each operating unit only every other year (stagger reviews as was done for Action Plans).
  - Guidelines, norms and oversight mechanisms should be developed for the technical review committees to help them work more objectively, consistently, and efficiently.
  - Milestones for providing feedback from the R4 reviews to operating units (scores, reporting cables, etc...) should be set that cover:
    - timeliness,
    - consistency, and
    - establishing points of contact.
  - Norms should be established with operating units that cover requests for information, sharing information on results of scoring, operating unit participation in the review process, and feedback expectations (a corollary to the management contract).
  - Every operating unit should be represented at its program review.
  - Ways should be explored to increase senior management participation in the review process, particularly program days. Potential ideas include a one-hour summary at the end of the day, a two-hour wrap-up meeting once a week, and round tables with country desk officers.
  - Best practices should be shared with operating units on how to improve communication within the operating unit on the R4 preparation and review process, including sharing feedback and scores from USAID/Washington more widely with staff.



**APPENDIX A:**

**LIST OF RESPONDENTS AND KEY PERSONS INTERVIEWED**

**APPENDIX B:**

**INTERVIEW GUIDE FOR KEY INFORMANT INTERVIEWS AND  
FOCUS GROUP INTERVIEWS**

## GUIDE FOR KEY INFORMANT AND FOCUS GROUP INTERVIEWS

### 1) SATISFACTION OF OBJECTIVES FOR THE R4

Key Questions: Were the objectives for the FY 1999 R4 met?

- Were the Agency's objectives for the FY 1999 R4 review met?
- Were Africa Bureau's objectives for the FY 1999 R4 review met?
- Were operating unit objectives for preparing and internally reviewing the FY 1999 R4 met?

Probes - Further Questions:

- How successful was preparation and internal review of the R4 by operating units in terms of their own assessment and evaluation of progress toward results?
- How successful were the R4 documents in providing information to USAID/Washington on the impact of operating unit programs?
- How successful was the R4 preparation and Africa Bureau review process in making sure that information on progress toward results was used as a factor in making decisions regarding budget allocation for operating unit programs?

Key Questions: Was there clarity around the objectives/purposes of the R4 within the Africa Bureau? Did participants all have the same understanding of the objectives of the R4?

Key Questions: Was Agency guidance for the R4 clear, adequate, and timely? Was Africa Bureau supplementary information clear, adequate, and timely?

### 2) USE OF R4 INFORMATION FOR MANAGING FOR RESULTS

Key Questions: Have operating units used information from the R4 preparation and internal review process to make any program decisions (i.e. to manage for results)? If so, are there any examples?

Key Questions: Who has used or is using FY 1999 R4 information, and for what purposes? What kinds of decisions have been or are being influenced by R4 information?

Probes - Further Questions: Have you used information from the Africa Bureau FY 1999 R4s or R4 reviews in your work? If so, how?

### 3) EFFECTIVENESS OF AFRICA BUREAU R4 REVIEW PROCESS

Key Questions: What were the major benefits or strengths of the FY 1999 R4 review process for the Africa Bureau? What were the main drawbacks or weaknesses?

Key Questions: Was the Africa Bureau FY 1999 R4 review process as effective as possible? How could effectiveness be improved?

Probes - Further Questions:

- How successful was the Africa Bureau R4 review process in analyzing SOs involving more than one funding sector? Did the budget category-based analysis system permit effective assessment of synergies in combined-sector SOs? What were the strengths and weaknesses of the system in terms of assessing cross-sectoral SOs?
- Was the entire R4 review process used by the Africa Bureau clear to all participants (procedures, different elements of scoring, different review committees, etc...)? If not, which parts were not clear?
- Which aspects of the technical review committees worked best, and which worst? How could the committee system be improved?
- How successful was the Africa Bureau R4 review process as a mechanism for USAID/Washington to provide feedback to operating units on performance?
- Did operating units have adequate opportunity to participate in the Africa Bureau R4 review process?
- Did operating units receive information from technical review committees in time to respond to issues or questions raised by the committees?
  - How could communication and interaction between field operating units and USAID/Washington be improved in the R4 preparation and review process?
  
- Were the criteria used by the Africa Bureau to review and rank SOs acceptable? Do you have any suggestions to improve any of the criteria?
- How does the Africa Bureau review process compare to that of other geographic bureaus in terms of effectiveness? (to be asked only of people who participated in reviews in other geographic bureaus)

4) EFFICIENCY OF R4 PREPARATION AND REVIEW PROCESS

Key Questions: Was the Africa Bureau FY 1999 R4 review process as efficient as possible? How could efficiency be improved? Did it involve a reasonable level of effort that it would be feasible to repeat every year?

Probes - Further Questions:

- Was the overall level of effort in the entire process reasonable?
- Was the workload of the technical review committees reasonable?

Key Question: Was the level of effort expended on R4 preparation by operating units reasonable?

5) IS THERE ANYTHING ELSE YOU WOULD LIKE TO COMMENT ON?

**APPENDIX C:**  
**QUESTIONNAIRE FOR OPERATING UNIT SURVEY**