

The African Evaluation Guidelines: 2002

A checklist to assist in planning evaluations, negotiating clear contracts, reviewing progress and ensuring adequate completion of an evaluation.

Nairobi M&E Network, African Evaluation Association Secretariat, Réseau Nigérien de Suivi et Evaluation, Cape Verde Evaluation Network, Réseau Malagache de Suivi et Evaluation, Comoros Evaluation Network, Eritrean Evaluation Network, Malawi M&E Network, Réseau National de Chercheurs et Evaluateurs de Burundi, Rwanda Evaluation Network, UNICEF Eastern and Southern Africa Region M&E Network¹

Abstract

A review of the US “*Program Evaluation Standards*”, undertaken in a series of workshops and meetings of networks of evaluators in Africa, resulted in modifications to those standards. The result was presented to a plenary session of the Inaugural Conference of the African Evaluation Association in September 1999, attended by over 300 evaluators from 35 countries. The AfrEA Conference decided that a systematic effort should be made to produce a list of African Evaluation Guidelines, similar to the Program Evaluation Standards, and that this checklist should be reviewed by national evaluation associations and networks in Africa and field tested in several countries. Eleven national and regional networks and associations suggested modifications to the text and endorsed the final version of the Guidelines.

Keywords

HIV/AIDS, African evaluation guidelines, programme evaluation, meta-evaluation, quality control

Introduction

A project to develop professional standards for program evaluation started in the USA in 1975. Its goal was to improve the quality of evaluations of educational and training programs, projects and materials. A Committee was formed, composed of 16 professional education associations and including the American Evaluation Association and a series of national professional associations interested in ensuring high quality in programme evaluations.² The Committee identified 30 aspects of good quality

¹ Mahesh Patel, Regional Monitoring and Evaluation Officer of the Eastern and Southern Africa Regional Office of UNICEF co-ordinated this process, in his capacity as Convenor of the African Evaluation Association. Corresponding author: Mahesh Patel, UNICEF ESARO, POB 44145, Nairobi, Kenya. email: mpatel@unicef.org

² The Professional Associations that are members of the Committee are: The American Association of School Administrators; American Educational Research Association; American Federation of Teachers; American Psychological Association; Association for Assessment in Counseling; Association for Supervision and Curriculum Development; Canadian Society for the Study of Evaluation; Council of Chief State School Officers; Council on Post Secondary Accreditation; National Association of Elementary

evaluations and presented them in the form of a 'checklist'. This checklist can be used to assess completeness and quality of evaluation work. The Committee produced two books on how to use these standards (JCSEE, 1989 & 1994). In 1989, the "*Program Evaluation Standards*" (PES) were approved by the American National Standards Institute (ANSI).

Individual evaluators can use the standards as a checklist while doing an evaluation, to ensure that most pertinent issues have been covered. It is also possible to use the checklist when negotiating a contract to do an evaluation, so that both parties have a clear understanding and agreement about what will be covered by the contract and what will not.

The 30 criteria in the checklist are now routinely used in the USA. Their usage in other countries is increasingly common. In Africa, Austral-Asia, Germany, South Korea and Switzerland, professional associations of evaluators are working to adapt the Program Evaluation Standards to make them more relevant to their national conditions.

Some donors have used the US PES to assess the quality of project evaluations performed in developing countries (Forss & Carlson, 1997). But not all of these US based 'Standards' are ideal for Africa. In some cases they are culturally inappropriate or misleading. This paper describes the process used to produce a set of African Evaluation Guidelines. The goal was to produce a set of Guidelines that are adopted by African evaluators and governments for evaluation of development programmes in Africa. The 2002 (fourth) version of the African Evaluation Guidelines is presented as an annex to this paper and is its first formal publication.

The national evaluation associations involved in the creation of this document recognize that these Guidelines are only a starting point, a dynamic work in progress. The process of consultation between national associations and further development of the African Evaluation Guidelines will continue and even intensify in the future.

Background

During the 1998 UNICEF Regional Evaluation Workshop in Nairobi, a training session was held on the 'Program Evaluation Standards', and a focus group discussion led by African evaluators was conducted on the theme "*Are the US 'Program Evaluation Standards' appropriate for use in African cultures?*" This discussion was followed by a visualized evaluation session on the same topic. A paper describing the proposed modifications to the US PES was produced (Patel & Russon, 1998). Later that year, the Nairobi Monitoring and Evaluation Network discussed this initial draft of modifications to the PES and suggested further changes.³ The revised draft was presented to and discussed by a group of young researchers in the Kenya Graduate Employment Programme and further amended.

School Principals; National Council for on Measurement in Education; National Education Association; and National School Boards Association.

³ . The Nairobi Monitoring and Evaluation Network later became the Kenya Evaluation Association.

The results of these discussions were presented to a plenary session of the Inaugural Conference of the African Evaluation Association (September 1999), with participation of over 300 African evaluators, as a draft document. Further modifications were suggested. Follow-up discussions were undertaken at the joint World Bank, African Development Bank and South African Development Bank “Regional Workshop on Monitoring and Evaluation Capacity Development in Africa” (September, 2000) and at a UNICEF Eastern and Southern Africa training workshop on “Monitoring and Evaluation in Emergency Situations” (February 2002).

The 1999 African Evaluation Association Conference recommended that these African Evaluation Guidelines should be reviewed by national networks and associations and field tested in Africa. The same year, field-testing was initially undertaken in two country evaluations (Zambia and Kenya) and in a multi-country evaluation. These results were favorable and further tests were undertaken in 2000 and 2001. By 2002, it became common for UNICEF funded evaluations in Eastern and Southern Africa to include the AEG as part of the protocol of ‘Terms of Reference’ of an evaluation. To date, the Guidelines have been used in ten countries in Africa.

Over the period 1999-2002, eleven networks and associations of evaluators in Africa reviewed the AEG. All these groups suggested modifications. These eleven groups will present this consolidated fourth version of the text to the Second African Evaluation Association Conference, in Nairobi in June 2002, for review and formal approval.

Results

This section synthesizes the key issues that emerged from the various workshops and conference based discussions and field tests that took place, and from the formal reviews by the national evaluation associations listed as joint authors of the Guidelines.

The Need for Guidelines

All the groups agreed that it is useful to have a set of quality enhancing guidelines for program evaluation research. The main reason given was the need to improve the quality of evaluative work. Government and donor agency concerns about program efficacy were often mentioned. The young researchers in the Kenya Graduate Employment Programme were particularly interested in having a checklist that could be used during contract negotiations. They wanted to have a good description of what should be covered in the evaluation and what could be omitted. Examples of contracting agencies that added tasks after completion of the report were mentioned.

Adoption or Adaptation?

While there was a strong consensus on the desirability of having guidelines, there was a lot of discussion over the types of guidelines that should be adopted. In the

early stages of this initiative, in 1999, there appeared to be three, somewhat overlapping, perspectives.

The first perspective was that it is acceptable to adopt an international model that had sufficient sensitivity to the African context. Many participants thought that the US Program Evaluation Standards (US PES) did not come laden with values that were in conflict with African values. This group said that there were no major cultural barriers to the use of the US Program Evaluation Standards in African countries in their original form.

The second perspective was that it was unacceptable to impose an externally developed set of standards on Africa. Proponents of this view thought that Africa should not 'submit' to a set of standards for which they had not provided any input. They felt that the US PES should either be rewritten with input from African stakeholders, or that African evaluators should develop their own standards. Subgroups considered either that a set of Africa level guidelines could be created that allowed local flexibility in their interpretation, or that each country (and perhaps by implication, institution or agency) should create their own guidelines.

The third perspective was that the appropriate procedure would be to test the US PES in field conditions in Africa in order to determine their suitability and to identify any modifications that might be required on a pan-African basis.

From 2000 onwards, a consensus was achieved that a checklist that would be useful for evaluators in Africa could be created through a consultative process and field-testing. This does not exclude further national modifications – indeed even the act of translation into local languages, as has been done in Burundi - is likely to generate local modifications.

Structure of the African Evaluation Guidelines

The 30 US PES are clustered in four groups that correspond to the four key attributes of sound and fair program evaluation. Some participants in one discussion thought there should be a fifth category, covering socio-cultural standards. The possibility of creating a fifth category was not followed-up, as it did not arise again in subsequent discussions.

Most comments focused on modifications of the existing groups, as detailed below. The 'Réseau Nigérian de Suivi et Evaluation' suggested that the definition of the Utility Category should be revised and this was done. The four categories used for the African Evaluation Guidelines, with the revised text in italics, are:

Utility - the utility guidelines are intended to help to ensure that an evaluation will serve the information needs of intended users *and be owned by stakeholders*

Feasibility - the feasibility guidelines are intended to help to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal

Propriety - the propriety guidelines are intended to help to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results

Accuracy - the accuracy guidelines are intended to help to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.

The Utility Guidelines

Many groups noted that social and cultural differences created some significant problems for the application of the US-PES “Utility Standards” in Africa. Modifications were proposed that resulted in more appropriate African formulations.

U1-PES Stakeholder Identification

“Persons involved in or affected by the evaluation should be identified, so that their needs can be addressed.”

It was agreed that the views of all stakeholders should be considered. Some noted that a key group of stakeholders, the program beneficiaries, often do not have organizations to represent them. Access to some geographical, ethnic or linguistic groups may be difficult for logistical or security reasons. Communications infrastructure is often not well developed. In developing country conditions, administrative infrastructure often does not extend far beyond the paved roads. Stakeholders within reach of paved roads are often over-sampled, one of the characteristics of ‘development tourism’. The Niger group considered that ownership was also an important issue. While the essence of the text was considered useful, some modification of the wording was considered desirable. The revised text, with the modified text in italics, is presented below.

U1-AEG “Stakeholder Identification. Persons and organizations involved in or affected by the evaluation (with special attention to beneficiaries at community level) should be identified and included in the evaluation process, so that their needs can be addressed and so that the evaluation findings are utilizable and owned by stakeholders, to the extent that this is useful, feasible and allowed.”

U4-PES Values Identification

“The perspectives procedures, and rationale used to interpret the findings should be carefully described, so that the bases for value judgments are clear.”

Discussions noted that the US-PES U4 does not specify which value system should be employed. The generic nature of the standard led many to consider that African values could be used just as easily as American or European values.

Others thought that different groups of stakeholders, such as the country governments, donors, implementing agencies and the evaluators, might come from different cultures and have different values. It is essential that evaluation methods preserve transparency in this area. In Africa, cultural diversity amongst stakeholders may be greater than in the

US. More than one value system may be in operation at the same time. Many reviewers thought that this standard should be strengthened. The revised text is presented below with the modifications proposed for the guideline in italics.

U4-AEG “Values Identification. The perspectives, procedures, and rationale used to interpret the findings should be carefully described, so that the bases for value judgements are clear. *The possibility of allowing multiple interpretations of findings should be transparently preserved, provided that these interpretations respond to stakeholders’ concerns and needs for utilization purposes.*”

U6-PES Report Timeliness and Dissemination

“Significant interim findings and evaluation reports should be disseminated to intended users, so that they can be used in a timely fashion.”

Discussions often raised the issue that concepts of time, perceptions of the importance of being “on time” and definitions of what “on time” actually means, are different in Africa. In Africa, the “way in which a thing is done” is often considered more important than getting it done “on time and within the budget”. Timeliness was also a standard that Smith et. al. (1993) thought should be modified to make the US-PES more relevant to Malta and India, noting that, *“To insist on holding someone to an officially stated deadline is viewed as nit-picking and unreasonable”*.

The injunction to disseminate reports to intended users created some discussion on the topic of, “Who are the intended users?” This issue was covered in the discussion of “P6. Disclosure of findings”, but it was proposed that the text of this Guideline also take these issues into account.

U6-AEG “Report Timeliness and Dissemination. Significant interim findings and evaluation reports should be disseminated to intended users, so that they can be used in a *reasonably* timely fashion, *to the extent that this is useful, feasible and allowed. Comments and feedback of intended users on interim findings should be taken into consideration prior to the production of the final report.*”

The Feasibility Guidelines

The Feasibility Guidelines attracted the largest number of comments. Feedback stated that cultural issues presented significant challenges in the application of the feasibility guideline to African contexts.

F2-PES Political Viability

“The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted.”

Some groups considered that this guideline presented special challenges to evaluators employed by implementing and donor agencies. There may be differences between the perspectives of donor groups and national governments. The former often have a significantly greater interest in impact evaluation than the latter. Project officers are typically more interested in using evaluation to improve implementation.

Within countries, cultures are often not homogeneous. In a supplementary challenge to political viability, civil conflict between different groups is not uncommon. Participation of some groups may not be politically feasible or consistent with security of the evaluator(s). For these or other reasons, governments and other agencies may wish to limit the participation of some groups. This guideline was relaxed slightly.

F2-AEG “Political Viability. The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted *to the extent that this is feasible in the given institutional and national situation.*”

F3-PES Cost Effectiveness

“The evaluation should be efficient and produce information of sufficient value, so that the resources expended can be justified.”

An occasional tension between program implementation and program evaluation comes from cost issues. Program officers are often keen to use almost all of their funding for programme implementation, rather than spend much of it on evaluation. On the other hand, donors may require formal evaluation, even if evaluation is considered an optional expense by governments or other stakeholders.

Evaluation budgets could certainly be more carefully estimated and actual expenditures on the evaluation more carefully monitored. The problem of cost over-runs during evaluation studies came up in several discussions. Several evaluators expressed the view that budgets should be monitored more carefully and that total expenditures should stay within budget. Consequently, the text of the African Evaluation Guideline includes the issue of monitoring of expenditures on evaluation and keeping within a budget.

F3-AEG “Cost Effectiveness. The evaluation should be efficient and produce information of sufficient value, so that the resources expended can be justified. *It should keep within its budget and account for its own expenditures.*”

The Propriety Guidelines

The propriety guidelines presented the greatest challenges of all the four categories of guidelines. The reason was the subjective nature of propriety. What is considered appropriate in one context and culture may be considered a serious error in another. Different cultures have different sets of values.

P2-PES Formal Agreements

“Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that these parties are obligated adhere to all conditions of the agreement or formally to renegotiate it.”

Many participants in discussions considered that the rule of law, development of formal agreements based upon law, and use of renegotiation and litigation in response to contract violations are quite different in Africa from in the US. Law is not always more important than tradition or custom. A formal agreement may not be honored if it contravenes custom or traditional law. Smith et. al. (1993) came to a similar conclusion, asserting that formal agreements, not related to property and tenancy, are uncommon and often unenforceable.

As in the discussion of timeliness, informal obligations and expectations are often more significant than those which are formally expressed. As development agencies are themselves composed of members from various cultures and may be contracting evaluation services from with members of yet another culture, there is considerable scope for confusion over informal expectations and interpretations of formal contractual arrangements. It is not always possible to resolve this confusion, or dispel culturally based implicit expectations through the medium of a written contract. Extended and repeated informal dialogue may often be more productive. The young researchers of the Kenya Graduate Employment Programme were particularly interested in the possibility of using the guidelines to make the expectations of employers as clear as possible and indicated that sometimes the informal expectations of agencies surprised them.

P2-AEG “Formal Agreements. Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to through dialogue and in writing, to the extent that this is feasible and appropriate, so that these parties have a common understanding of all the conditions of the agreement and hence are in a position to formally renegotiate it if necessary. Specific attention should be paid to informal and implicit aspects of expectations of all parties to the contract.”

P3-PES Rights of Human Subjects

“Evaluation should be designed and conducted to respect and protect the rights and welfare of human subjects.”

It was commonly noted that, in developed countries, human rights tend to focus upon the rights of the individual. In Africa, and indeed in developing countries in other regions, people tend to have stronger ties to the extended family, tribe and other groups than do people in developed countries. In these interdependent social systems, individual rights are balanced by obligations. The rights of the individual are perhaps more often considered in relation to, balanced by, or even in some instances subordinate to, the welfare of the community. The US-PES does not make any useful distinction between individual and collective rights. In Africa, notions of collective rights of communities are

much more developed, covering even rights to land in many cultures. The African Evaluation Guideline on “Rights of human subjects” takes account of these more extensive concepts of rights and pays explicit attention to the issue of confidentiality of personal information.

P3-AEG “Rights of Human Participants. Evaluation should be designed and conducted to respect and protect the rights and welfare of human subjects and the communities of which they are members. The confidentiality of personal information collected from various sources must be strictly protected.”

P4-PES Human Interaction

“Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation, so that participants are not threatened or harmed.”

In some cultures there are very strong limits to the type of interactions that an evaluator can have with the persons associated with the evaluation. These can be especially important when stakeholders from one culture are interviewing stakeholders from another. There may be prohibitions against interactions between genders and even between population groups. Some cultural and religious systems place specific limits on permissible interactions between men and women, and for some topics, also on dialogue between members of the same sex. The range of culturally acceptable interaction is likely to vary with culture and religion.

P4-AEG “Human Interaction. Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation, so that participants are not threatened or harmed or their cultural or religious values compromised.”

P6-PES Disclosure of Findings

“The formal parties to an evaluation should ensure that the full set of evaluation finding along with pertinent limitations are made accessible to the persons affected by the evaluation, and any others with expressed legal rights to receive the results.”

There were often some tensions during the discussion of this PES. Many felt that this PES was only applicable in a democratic country. They considered that “P6” was not applicable in countries that are subject to dictatorship or authoritarian government. In that political setting, communication of findings can neither be assumed nor assured in advance. In extreme cases, even a request by an evaluator to assure the eventual release of findings could be viewed with suspicion. The release of findings without government approval in such settings could well be dangerous, especially for national evaluators, and very rarely occurs. International journalists are a special case here and it is not uncommon that they may be declared ‘persona non grata’ and asked to leave the country.

In other settings, this may not be a significant issue. Indeed, evaluators from countries with more liberal governments would often argue that “P6” should not be “watered down” in order to encourage authoritarian governments to liberalize their practices. More seriously, some considered that any relaxation of this guideline would provide a loophole that positively encouraged ‘in-between’ governments to react in a conservative manner.

It was also recognized that development agencies, in general, reserved copyright over the findings of the evaluations that they financed and that national evaluators normally did not have the leverage to revise the way in which these agencies normally work.

A consensus on the formulation below was achieved in order to maintain hegemony amongst countries, while protecting those evaluators in Africa who are required to work with difficult regimes as well as those required to respect the copyright of commissioning agencies.

P6-AEG “Disclosure of Findings. The formal parties to an evaluation should ensure that the full set of evaluation findings along with pertinent limitations are made accessible to the persons affected by the evaluation, and any others with expressed legal rights to receive the results as far as possible. The evaluation team and the evaluating institution will determine what is deemed possible, to ensure that the needs for confidentiality of national or governmental entities and of the contracting agents are respected, and that the evaluators are not exposed to potential harm.”

P7-PES Conflict of Interest

“Conflict of interest should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results.”

The notion of conflict of interest has some special dimensions in Africa. One participant noted that an evaluation group that had been awarded the contract to evaluate the entire programme of a United Nations agency was also competing for additional United Nations contracts and grants. The evaluation group did not see this as a conflict of interest. (If they were not awarded subsequent contracts would their report portray the UN agency unfavorably? Alternatively, if the evaluation was critical, might the group not be awarded further contracts?) In the cited case, the contractors thought that they would be better qualified to undertake additional contracts because of their enhanced familiarity with the programming modalities of that agency.

However, even if the items that might create conflict are different, the notion of resolving possible and even actual conflict through discussion between the parties involved is very African. While this was much discussed, no modification to the Guidelines was proposed.

The Accuracy Guidelines

At a general level, some discussants noted that accuracy might sometimes be compromised by cultural factors. In Africa, an evaluative report might tend to be diplomatically supportive in a selective manner, rather than comprehensive and critical.

A1-PES Program Documentation

“The program being evaluated should be described and documented clearly and accurately, so that the program is clearly identified.”

In some settings, particularly in rural areas, documentation may be rather scarce. Communities do have information, even if their members are mostly illiterate, and in these cases descriptions would have to be elicited verbally. There was a clear consensus that evaluators would need to pay special attention to oral histories and traditional modes of recording information.

A1-AEG “Program Documentation. The program being evaluated should be described clearly and accurately, so that the program is clearly identified, with attention paid to personal and verbal communications as well as written records.”

A2-PES Context Analysis

“The context in which the program exists should be examined in enough detail, so that its likely influences on the program can be identified and assessed.”

Context analysis was frequently discussed. Most groups concluded that context analysis, like the Guideline dealing with value identification, is one of that is already oriented towards international and cross-cultural applications. It was noted that many of the subtle nuances of the evaluation would be captured in a context analysis if it were comprehensive. Context analysis could reveal possible bias and perhaps evidence of situations where certain methods or approaches might not yield positive results. One national network considered that several specific aspects of context analysis should explicitly be stated and their proposed amendment was accepted.

A2-AEG “Context Analysis. The context in which the program exists should be examined in enough detail, including political, social, cultural and environmental aspects, so that its likely influences on the program can be identified and assessed.”

A4-PES Defensible Information Sources

“The sources of information used in a program evaluation should be described in enough detail, so that the adequacy of the information can be assessed.”

Some religious systems in Africa prohibit various forms of contact between women and men, especially with outsiders. Discussions on sensitive topics such as reproductive health or HIV/AIDS are strictly taboo in some settings. A male evaluator may not be permitted to administer a questionnaire on sexual practices directly to a woman. At the same time, a female evaluator might not be admitted to the household at all. If the husband considered that women should not leave the home, he might consider the female

evaluator to be setting a bad example to other members of the household. (Some additional implications of this are explored in the discussion of the next standard.)

While it is important to identify the types of sources used, in some cases it may be especially important to protect anonymity of individual sources and even to avoid the possibility of identifying specific communities.

The pursuit of ‘adequacy of the information’ requires a special sensitivity to cultural and political considerations in countries.

A4-AEG “Defensible Information Sources. The sources of information used in a program evaluation should be described in enough detail, so that the adequacy of the information can be assessed, without compromising any necessary anonymity or cultural or individual sensitivities of respondents.”

A5-PES Valid Information

“The information gathering procedures should be chosen or developed and then implemented so that they will assure that the interpretation arrived at is valid for the intended use.”

An evaluator who is required to administer a questionnaire on gender specific aspects of behavior to a woman through her father or husband may not receive valid information. The father or husband may not have accurate knowledge of the women’s behavior and may not admit that but instead give inaccurate answers to the questions.

Some information sources may fear answering questions accurately because the answers would contradict an official government position or expose them to other risks. Questions answered under duress may not be valid or reliable.

Finally, evaluators posing questions to beneficiary groups are often closely identified with the donors whose programme is being evaluated. It is traditional, in many parts of Africa, to attempt to anticipate desired responses and provide answers that would reflect positively on the programme.

A5-AEG “Valid Information. The information gathering procedures should be chosen or developed and then implemented so that they will assure that the implementation arrived at is valid for the intended use. Information that is likely to be susceptible to biased reporting should be checked using a range of methods and from a variety of sources.”

A9-PES Analysis of Qualitative Information

“Qualitative information in an evaluation should be appropriately and systematically analyzed so that evaluation questions are effectively answered.”

This Guideline was often discussed though no modification was proposed. Quite a strong consensus opinion was expressed that African traditions of passing knowledge down by

word of mouth lend themselves more readily to qualitative evaluation methods than to quantitative methods.

Conclusions

The US Program Evaluation Standards were found to be a useful starting point for the development of African Evaluation Guidelines. The PES were changed to make them more readily applicable to current African cultural, social and political realities. Of the original 30 US PES, 13 have so far been revised and 17 remain unchanged. Political and cultural considerations emerged as major driving forces behind the necessary modifications.

Guidelines on “Political Viability” (F2) and “Disclosure of Findings” (P6) were both considered politically sensitive in some countries – but not necessarily in all. The wording of the guidelines is a compromise between the proposals of countries with relatively open governments, freedom of press and generally participative political processes and those with relatively autocratic governments or military dictatorships.

Cultural considerations were important in the wording of several guidelines, especially those relating to propriety. “Formal Agreements” (P2), “Rights of Human Subjects” (P3), and “Human Interactions” (P4) all required modification, as did the one on “Defensible Information Sources” (A4). The Guideline on “Valid Information” (A5) was adjusted in consideration of cultural sensitivities to permitted male-female interactions and to queries on topics such as sexual behavior. One example is the Guideline on “Stakeholder identification” (U1), which was extended to pay explicit attention to the sometimes-ignored beneficiaries at community level.

Two supplementary issues are worthy of consideration here. Firstly, the English language itself is often considered to contain implicit cultural concepts and assumptions. The African Evaluation Guidelines have been translated into French in Burundi. It is anticipated that further translations into local languages will yield additional insights and modifications.

Secondly, in the US the PES are published as a book that contains examples of the practical use of the PES in a variety of settings. The use of these examples from evaluation practice allows a more in-depth understanding of their meaning and utility. These examples are of course based on the practice of evaluation in the USA. For the PES to fulfill their full potential in Africa it may be useful to have a similar book length treatment of the subject, using concepts and examples derived from evaluation practice in Africa.

The African Evaluation Association Working Group on this subject will seek donor funding in support of this effort as donor agencies have often expressed interest in helping to enhance the quality of evaluation work in Africa.

We hope that the list of African Evaluation Guidelines presented in Annex 1 will be useful to those wishing to use a quality control checklist when performing or reviewing evaluations. To facilitate that process, Annexes 2, 3, and 4 present the Guidelines in matrix form, as they would be used in practice. Not all Guidelines will be relevant for all evaluations. The matrix presentation allows the evaluator and the funding agency to decide together which ones are applicable. This may help to avoid implicit assumptions about the scope of the evaluation from surfacing after the evaluation is already complete. The matrices function as an effective evaluation management tool, or checklist, that can be used by the evaluator alone, or in collaboration with the contractor.

Cumulative experience in Africa in working with the Guidelines will further enhance their utility and may result in improvements. The African Evaluation Association Secretariat requests all evaluators in Africa to send in evaluations that make use of these Guidelines as a means of improving the Guidelines and for inclusion in a book on this subject.

Bibliography

Forss K & Carlsson J. The quest for quality – Or can evaluation findings be trusted? Evaluation (1997) V.3, N.4, 481-501.

Joint Committee on Standards for Educational Evaluation (1981) Standards for evaluations of educational programmes, projects and materials. New York, McGraw-Hill.

Joint Committee on Standards for Educational Evaluation (1988) The Personnel Evaluation Standards. Newbury Park, CA: Sage

Joint Committee on Standards for Educational Evaluation. (1994). The Program Evaluation Standards. Sage: Thousand Oakes, CA

Patel M. & Russon C. Appropriateness of the Program Evaluation Standards for use in Africa. 1998. Presented in modified form to the African Evaluation Association in 1999.

Smith N., Chircop S., & Mukherjee P., Considerations on the development of culturally relevant evaluation standards. Studies in Educational Evaluation (1993) V.19, 3-13.

Annex 1: The African Evaluation Guidelines: 2002

A checklist to assist in planning evaluations, negotiating clear contracts, reviewing progress and ensuring adequate completion of an evaluation.

Nairobi M&E Network, African Evaluation Association Secretariat, Réseau Nigérian de Suivi et Evaluation, Cape Verde Evaluation Network, Réseau Malagache de Suivi et Evaluation, Comoros Evaluation Network, Eritrean Evaluation Network, Malawi M&E Network, Réseau National de Chercheurs et Evaluateurs de Burundi, Rwanda Evaluation Network, UNICEF Eastern and Southern Africa Region M&E Network

Utility: The utility guidelines are intended to ensure that an evaluation will serve the information needs of intended users *and be owned by stakeholders*.

U1. (modified) Stakeholder Identification. Persons and organizations involved in or affected by the evaluation (with special attention to beneficiaries at community level) should be identified and included in the evaluation process, so that their needs can be addressed and so that the evaluation findings are utilizable and owned by stakeholders, to the extent this is useful, feasible and allowed.

U2 Evaluator Credibility. The persons conducting the evaluation should be both trustworthy and competent to perform the evaluation, so that the evaluation findings achieve maximum credibility and acceptance.

U3 Information Scope and Selection. Information collected should be broadly selected to address pertinent questions about the program and be responsive to the needs and interests of clients and other specified stakeholders.

U4 (modified) Values Identification. The perspectives, procedures, and rationale used to interpret the findings should be carefully described, so that the bases for value judgments are clear. The possibility of allowing multiple interpretations of findings should be transparently preserved, provided that these interpretations respond to stakeholders' concerns and needs for utilization purposes.

U5 Report Clarity. Evaluation reports should clearly describe the program being evaluated, including its context, and the purposes, procedures, and findings of the evaluation, so that essential information is provided and easily understood.

U6 (modified) Report Timeliness and Dissemination. Significant interim findings and evaluation reports should be disseminated to intended users, so that they can be used in a reasonably timely fashion, to the extent that this is useful, feasible and allowed. Comments and feedback of intended users on interim findings should be taken into consideration prior to the production of the final report.

U7 Evaluation Impact. Evaluations should be planned, conducted, and reported in ways that encourage follow through by stakeholders, so that the likelihood that the evaluation will be used is increased.

Feasibility: The feasibility guidelines are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.

F1 Practical Procedures. The evaluation procedures should be practical, to deep disruption to a minimum while needed information is obtained.

F2 (modified) Political Viability. The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted to the extent that this is feasible in the given institutional and national situation.

F3 (modified) Cost Effectiveness. The evaluation should be efficient and produce information of sufficient value, so that the resources expended can be justified. It should keep within its budget and account for its own expenditures.

Propriety - The propriety guidelines are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.

P1 Service Orientation. Evaluation should be designed to assist organizations to address and effectively serve the needs of the full range of targeted participants.

P2 (modified) Formal Agreements. Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to through dialogue and in writing, to the extent that this is feasible and appropriate, so that these parties have a common understanding of all the conditions of the agreement and hence are in a position to formally renegotiate it if necessary. Specific attention should be paid to informal and implicit aspects of expectations of all parties to the contract.

P3 (modified) Rights of Human Participants. Evaluation should be designed and conducted to respect and protect the rights and welfare of human subjects and the communities of which they are members. The confidentiality of personal information collected from various sources must be strictly protected.

P4 (modified) Human Interaction. Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation, so that participants are not threatened or harmed or their cultural or religious values compromised.

P5 Complete and Fair Assessment. The evaluation should be complete and fair in its examination and recording of strengths and weaknesses of the program being evaluated, so that strengths can be built upon and problem areas addressed.

P6 (modified) Disclosure of Findings. The formal parties to an evaluation should ensure that the full set of evaluation finding along with pertinent limitations are made accessible to the persons affected by the evaluation, and any others with expressed legal rights to receive the results as far as possible. The evaluation team and the evaluating institution will determine what is deemed possible, to ensure that the needs for confidentiality of national or governmental entities and of the contracting agents are respected, and that the evaluators are not exposed to potential harm.

P7 Conflict of Interest. Conflict of interest should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results.

P8 Fiscal Responsibility. The evaluator's allocation and expenditure of resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible, so that expenditures are accounted for and appropriate.

Accuracy - The accuracy guidelines are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth of merit of the program being evaluated.

A1 (modified) Program Documentation. The program being evaluated should be described clearly and accurately, so that the program is clearly identified, with attention paid to personal and verbal communications as well as written records.

A2 (modified) Context Analysis. The context in which the program exists should be examined in enough detail, including political, social, cultural and environmental aspects, so that its likely influences on the program can be identified and assessed.

A3 Described Purposes and Procedures. The purposes and procedures of the evaluation should be monitored and described in enough detail, so that they can be identified and assessed.

A4 (modified) Defensible Information Sources. The sources of information used in a program evaluation should be described in enough detail, so that the adequacy of the information can be assessed, without compromising any necessary anonymity or cultural or individual sensitivities of respondents.

A5 (modified) Valid Information. The information gathering procedures should be chosen or developed and then implemented so that they will assure that the implementation arrived at is valid for the intended use. Information that is likely to be susceptible to biased reporting should be checked using a range of methods and from a variety of sources.

A6 Reliable Information. The information gathering procedures should be chosen or developed and then implemented so that they will assure that the information obtained is sufficiently reliable for the intended use.

A7 Systematic Information. The information collected, processed, and reported in an evaluation should be systematically reviewed and any errors found should be corrected.

A8 Analysis of Quantitative Information. Quantitative information in an evaluation should be appropriately and systematically analyzed so that evaluation questions are effectively answered.

A9 Analysis of Qualitative Information. Qualitative information in an evaluation should be appropriately and systematically analyzed so that evaluation questions are effectively answered.

A10 Justified Conclusions. The conclusions reached in an evaluation should be explicitly justified, so that stakeholders can assess them.

A11 Impartial Reporting. Reporting procedures should guard against distortion caused by personal feelings and biases of any party to the evaluation, so that evaluation reports fairly reflect the evaluation findings.

A12 Meta-evaluation. The evaluation itself should be formatively and summatively evaluated against these and other pertinent guidelines, so that its conduct is appropriately guided and, on completion, stakeholders can closely examine its strengths and weakness.

Annex 2: African Evaluation Guidelines – Checklist for Discussion between Contractor and Evaluator at Start of Evaluation

| Guidelines | Relevant | Irrelevant | Details of Type of Coverage or Consideration of this Issue Required |
|---|-----------------|-------------------|--|
| UTILITY GUIDELINES | | | |
| U1 Stakeholder Identification | | | |
| U2 Evaluator Credibility | | | |
| U3 Information Scope and Selection | | | |
| U4 Values Identification | | | |
| U5 Report Clarity | | | |
| U6 Report Timeliness and Dissemination | | | |
| U7 Evaluation Impact | | | |
| FEASIBILITY GUIDELINES | | | |
| F1 Practical Procedures | | | |
| F2 Political Viability | | | |
| F3 Cost Effectiveness | | | |
| PROPRIETY GUIDELINES | | | |
| P1 Service Orientation | | | |
| P2 Formal Agreements | | | |
| P3 Rights of Human Participants | | | |
| P4 Human Interactions | | | |
| P5 Complete and Fair Assessment | | | |
| P6 Disclosure of Findings | | | |
| P7 Conflict of Interests | | | |
| P8 Fiscal Responsibility | | | |
| ACCURACY GUIDELINES | | | |
| A1 Program Documentation | | | |
| A2 Context Analysis | | | |
| A3 Described Purposes and Procedures | | | |
| A4 Defensible Information Sources | | | |
| A5 Valid Information | | | |
| A6 Reliable Information | | | |
| A7 Systematic Information | | | |
| A8 Analysis of Quantitative Information | | | |
| A9 Analysis of Qualitative Information | | | |
| A10 Justified Conclusions | | | |
| A11 Impartial Reporting | | | |
| A12 Metaevaluation | | | |

Annex 3: African Evaluation Guidelines – Matrix to Check Adequacy of the First Draft of an Evaluation

| Guidelines | Adequate | Inadequate | Irrelevant | Suggestions on How to Achieve Adequacy |
|---|----------|------------|------------|--|
| UTILITY GUIDELINES | | | | |
| U1 Stakeholder Identification | | | | |
| U2 Evaluator Credibility | | | | |
| U3 Information Scope and Selection | | | | |
| U4 Values Identification | | | | |
| U5 Report Clarity | | | | |
| U6 Report Timeliness and Dissemination | | | | |
| U7 Evaluation Impact | | | | |
| FEASIBILITY GUIDELINES | | | | |
| F1 Practical Procedures | | | | |
| F2 Political Viability | | | | |
| F3 Cost Effectiveness | | | | |
| PROPRIETY GUIDELINES | | | | |
| P1 Service Orientation | | | | |
| P2 Formal Agreements | | | | |
| P3 Rights of Human Participants | | | | |
| P4 Human Interactions | | | | |
| P5 Complete and Fair Assessment | | | | |
| P6 Disclosure of Findings | | | | |
| P7 Conflict of Interests | | | | |
| P8 Fiscal Responsibility | | | | |
| ACCURACY GUIDELINES | | | | |
| A1 Program Documentation | | | | |
| A2 Context Analysis | | | | |
| A3 Described Purposes and Procedures | | | | |
| A4 Defensible Information Sources | | | | |
| A5 Valid Information | | | | |
| A6 Reliable Information | | | | |
| A7 Systematic Information | | | | |
| A8 Analysis of Quantitative Information | | | | |
| A9 Analysis of Qualitative Information | | | | |
| A10 Justified Conclusions | | | | |
| A11 Impartial Reporting | | | | |
| A12 Metaevaluation | | | | |

Annex 4: African Evaluation Guidelines – Matrix to Check Adequacy of Completion of Contract/Work at the End of Evaluation

| Guidelines | Met | Partly Met | Not Met | Irrelevant | Suggestions for Ideas for improvement |
|---|-----|------------|---------|------------|---------------------------------------|
| UTILITY GUIDELINES | | | | | |
| U1 Stakeholder Identification | | | | | |
| U2 Evaluator Credibility | | | | | |
| U3 Information Scope and Selection | | | | | |
| U4 Values Identification | | | | | |
| U5 Report Clarity | | | | | |
| U6 Report Timeliness and Dissemination | | | | | |
| U7 Evaluation Impact | | | | | |
| FEASIBILITY GUIDELINES | | | | | |
| F1 Practical Procedures | | | | | |
| F2 Political Viability | | | | | |
| F3 Cost Effectiveness | | | | | |
| PROPRIETY GUIDELINES | | | | | |
| P1 Service Orientation | | | | | |
| P2 Formal Agreements | | | | | |
| P3 Rights of Human Participants | | | | | |
| P4 Human Interactions | | | | | |
| P5 Complete and Fair Assessment | | | | | |
| P6 Disclosure of Findings | | | | | |
| P7 Conflict of Interests | | | | | |
| P8 Fiscal Responsibility | | | | | |
| ACCURACY GUIDELINES | | | | | |
| A1 Program Documentation | | | | | |
| A2 Context Analysis | | | | | |
| A3 Described Purposes and Procedures | | | | | |
| A4 Defensible Information Sources | | | | | |
| A5 Valid Information | | | | | |
| A6 Reliable Information | | | | | |
| A7 Systematic Information | | | | | |
| A8 Analysis of Quantitative Information | | | | | |
| A9 Analysis of Qualitative Information | | | | | |
| A10 Justified Conclusions | | | | | |
| A11 Impartial Reporting | | | | | |
| A12 Metaevaluation | | | | | |

