The African Evaluation Guidelines 2000:

A checklist to assist evaluators: Draft 1.2

A joint draft working paper produced by, in chronological order, the:

Nairobi M&E Network
African Evaluation Association Secretariat
Réseau Nigérian de Suivi et Evaluation
Cape Verde Evaluation Network
Réseau Malagache de Suivi et Evaluation (Madagascar)
Eritrean Evaluation Network

Current Status: to be reviewed by - National Focal Points or Associations:

Comoros - Kamalidine Souef - ksouef@unicef.org

Ethiopia - Koorosh Raffii - kraffii@unicef.org

Ghana - Bishop Akolgo - isodec@ncs.com.gh

Kenya - Karen Odhiambo - karenodhiambo@hotmail.com

Rwanda - James Mugaju - jmugaju@unicef.org

South Africa - Edward Clarke - eclarke@netactive.co.za

Zambia - Martha Sichone - hodi@zamnet.zm

Zimbabwe- Mufani Khosa - mkhosa@mandel.delta.co.zw

The starting point for this initiative was the decision of the African Evaluation Association Conference to create a set of "African Evaluation Guidelines" based on a presentation to the plenary session on the subject of adapting the US "Program Evaluation Standards". It was decided to set up a committee to produce a draft document called the "African Evaluation Guidelines".The goal is now to produce a revised working draft based on comments from countries. That final draft document would be jointly authored by all national associations or networks that have participated in the review process and that wish to be associated with this initiative. National associations or networks may delegate the review to an appointed focal point, distribute this draft for comments, meet and discuss it as a group, or use some other process at their discretion. Unless otherwise informed, the AfrEA Secretariat will consider the leader of the national group to be the focal point for this activity.

Please provide feedback on whether you are interested in participating as a co-author by the end of this year, and on your conclusions by the end of January 2001. I am proposing a tight schedule as the Committee has not met during the last year and so this activity is now well behind schedule. It must be brought back on track for us to maintain momentum. If you have difficulties with this process or the schedule - and especially if you have a helpful proposal on how either can be improved - do let me know.

The African Evaluation Guidelines 2000:

A checklist to assist evaluators
Draft 1.4, Nairobi, January 2001

Abstract

A review of the relevance of the US "Program Evaluation Standards" (PES) to evaluation work in Africa was undertaken in a workshop setting, at the Inaugural Conference of the African Evaluation Association, in several meetings of the Kenya Evaluation Association and at a meeting on building evaluation capacity in Africa. Discussions in these venues showed that the US PES were a tool that could be used by evaluators in Africa as a checklist when planning evaluations, negotiating contracts and reviewing progress in implementation of an evaluation. However, it was also thought that several of the PES would need to be modified to make the list more appropriate for African conditions. The African Evaluation Association proposed that a set of "African Evaluation Guidelines" (AEG) be created, using the US PES as a starting point. The first draft of the African Evaluation Guidelines is attached as Annex 1.

In order to achieve consensus on the proposed "African Evaluation Guidelines" (AEG), a more structured and formal consultation process is now being undertaken. Existing national evaluation associations and networks will be offered the chance to review the current draft AEG and propose any additional modifications they

require. The revised document will then be disseminated under the joint authorship of the national associations that agree to the content of the joint document as an interim working paper .The publication will carry a date (2000, perhaps) or version number (1.0) as it is anticipated that future work by national associations in Africa may result in further development of the African Evaluation Guidelines. The provisional list of AEG is attached as Annex 1.

Introduction

A pioneer project to develop professional standards for programme evaluation was initiated in the USA in 1975.[1]Its goal was to improve the evaluation of educational and training programs, projects and materials in a variety of settings. A Joint Committee composed of 16 professional education associations and including the American Evaluation Association and the American Psychological Association was established. The Joint Committee compiled a set of 30 criteria, the "Programme Evaluation Standards" (PES). The PES took the form of a 'checklist' against which completeness and quality of evaluation work can be assessed. In 1989, the PES were approved by the American National Standards Institute (ANSI).

Individual evaluators can use the standards while doing an evaluation as a checklist to ensure that most pertinent issues have been covered. Performing such a 'meta-evaluation', an assessment of the evaluation itself, is acknowledged as good practice by the American Evaluation Association. 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.

These 30 criteria are now routinely used in the USA and their usage in other countries is increasingly common. In Australasia, Germany, South Korea and Switzerland, groups are working to adapt Program Evaluation Standards to make them more relevant to their national conditions.

Some donors have used the US PES to assess the quality of evaluations of projects implemented in developing countries.[2]It is not clear that these US 'Standards' can easily be applied 'in the field' in Africa. In some cases they may be culturally inappropriate or misleading. This paper is a step along the path of creating a set of African Evaluation Guidelines – not the final step, but a first step.

The national evaluation associations involved in the creation of this document consider that these Guidelines are a useful checklist to consider when planning an evaluation, negotiating a contract to do an evaluation and when reviewing progress during implementation of an evaluation. We further consider 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. The current version of the African Evaluation Guidelines described in this document is a snapshot of the evolution of that process at this point in time.

Background

During the 1998 UNICEF Evaluation Workshop in Nairobi, a training session was held on the 'Program Evaluation Standards'. The following day a focus group discussion 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 (VIPP) session on the same topic. Some modifications to the US-PES were proposed. Later in the year, the Nairobi Monitoring and Evaluation Network discussed the initial draft of modifications to the PES that emerged from the UNICEF workshop 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 this groups made several additional useful contributions.

The results of these discussions were presented to the Inaugural Conference of the African Evaluation Association (September 1999) as a draft document and some further modifications were suggested.[4] A follow-up discussion was also undertaken at the World Bank, African Development Bank and South African Development Bank Regional Workshop on monitoring and evaluation capacity development in Africa (September, 2000). The African Evaluation Association also requested that these African Evaluation Guidelines be field tested in Africa. This was undertaken in two country evaluations (Zambia and Kenya) and one multi-country evaluation in 1999.

The spirit of this enterprise has been to undertake all necessary changes to the PES, but where changes were not needed, to stay with the original formulation as much as possible. This approach was considered to maximize international comparability of results obtained and yet maintain the necessary African cultural sovereignty.

Results

The section synthesizes the issues that emerged from the various workshop and conference based discussions so far as the basis for a more formal review by the national evaluation associations listed on the front page.

The Need For Guidelines

Most discussions achieved a consensus that there was some utility to having a set of quality enhancing

guidelines for programme evaluation research. Reasons that were given included the need to improve quality of evaluative work. Government and donor agency concerns about programme efficacy were also often mentioned. The Graduate Employment Programme researchers were particularly interested in having available a list of items to them that could be used during contract negotiations to be sure to have a good description of what should be covered in the evaluation and what could be omitted. Instances of contracting agencies adding tasks after completion of the report were mentioned.

Adoption Or Adaption

While a consensus on the desirability of having guidelines was generally easily achieved, there was considerable discussion over the types of guidelines that should be adopted. There appeared to be three positions. The first was that it is acceptable to adopt an international model that had sufficient sensitivity to the African context. Many of the participants felt that the US Program Evaluation Standards did not come laden with values that were in conflict with African values. This group felt that, with the few modifications discussed below, there were no major cultural barriers to the use of the Program Evaluation Standards in African countries.

The second position was that it is unacceptable to impose an externally developed set of standards on Arica. 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.

Participants felt that the appropriate procedure would be to test the US PES in field conditions in Africa in order to determine their suitability and identify any modifications that might be required. This has indeed been done and one of the complications that was noted is that the situation in Africa is rendered more complex by the presence of donors and implementing agencies additional to government.

Structure of the African Evaluation Guidelines

30 US PES were categorized into four groups corresponding to the four attributes of sound and fair program evaluation identified by the members of the working group. Some participants in one discussion thought there should be a fifth category: socio-cultural standards. Many of the comments on the existing standards related to social and cultural issues. An alternative to modifying all or some of those so identified standards might be to create an additional category. This possibility was not followed in this work as it did not arise again. The 'Réseau Nigérian de Suivi et Evaluation' suggested that the definition of the Utility group 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

Considered that social and cultural differences created several significant problems for the application of the US-PES "Utility Standards" in Africa. Modifications were proposed that resulted in more appropriate formulations.

U1.Stakeholder Identification.

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

There was general agreement that one needs to take into consideration the views of all stakeholders. However, some contributors to discussion recognized that a key group of stakeholders, the program beneficiaries, often do not have organizations to represent them. Access to some geographical, ethnic of 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 tarmac. Stakeholders within reach of passable 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.(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 the evaluation findings can be operational and owned by stakeholders, to the extent this is useful, feasible and allowed.

U4.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.

This standard is an injunction to identify the perspectives, procedures, and rationale used to interpret the findings. 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.

Different groups of stakeholders, such as the country governments, donors, implementing agencies and the evaluators, may 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 rather greater than it is in the US. Many discussants considered that this standard should be strengthened. The revised text is presented below with the modifications proposed for the guideline in italics. The possibility of allowing multiple interpretations of findings was introduced to preserve major differences in perceptions that may occur. In an extreme case, program beneficiaries may consider that a program should be expanded while the donors consider that it should be closed. Differences in culture may lead to different perceptions and valuations of impact. Evaluators should preserve in the report these varied perceptions and the means by which conclusions are reached and indicate in the report the perceived basis of differences and their implications for 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.

U6.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. [5] They stated, "To insist on holding someone to an officially stated deadline is viewed as nit-picking and unreasonable".

In discussions in Africa two points of view were raised. One that time was perceived differently than in the developed world. The other that Africa should aspire to greater timeliness. It is useful to recall at this point that these guidelines are intended only as a checklist of issues that should be considered in negotiating and implementing an evaluation. The parties concerned would have to arrive at their own agreed interpretation. The purpose of the guidelines is to raise the issue for discussion in an appropriate manner.

The injunction to disseminate reports to intended users caused participants to raise the question, "Who are the intended users?" This issue was also covered in the discussion surrounding "P6. Disclosure of findings", but it was also proposed that the text of this Guideline take these issues into account.

It was noted that dissemination of interim findings was not the end point of the feedback and 'course correction' process. It was also essential that the comments of users of the evaluation be taken into account prior to the production of the final report.

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.

The Feasibility Guidelines.

The Feasibility Guidelines attracted the largest number of comments. Discussions noted that cultural issues presented significant challenges in the application of the feasibility criteria to African contexts.

F2. 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 discussants considered that this standard presented special challenges to evaluators employed by implementing donor agencies. Evaluation staff in these agencies are typically hierarchically subordinate to the Program Officers whose activities are to be evaluated. Programme Officers are sometimes over-ambitious in defining the objectives of a project and this may lead them to tend to omit, limit, or seek to control evaluation activities that might result in adverse reflections on their performance. There may also be differences between the perspectives of donor groups and national governments. The former often have a significantly greater interest in evaluation (especially of 'impact' of donor funded projects) than the latter.

Within countries, cultures are often not homogeneous. 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.

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 to counteracted to the extent that this is feasible in the given institutional and national situation.

F3. Cost Effectiveness.

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

Some part of an occasional tension between program implementation and evaluation activities comes from cost issues. Program officers are often keen to use all of their funding for programme implementation, rather than spend any of it on evaluation. On the other hand, donors may require formal evaluation, even if it is considered an unnecessary expense by other stakeholders such as the government.

Development assistance is often 'high risk' capital. It may face uncertain returns in its efforts to relieve human suffering in operationally difficult areas. It may face uncertain returns in its quest for useful innovations. Project goals may be formulated with insufficient advance knowledge of field conditions in that area. Given the sometimes high cost of quality information, errors such as these are perhaps inevitable. The value of formally structured evaluative information is high in principle, but very uncertain in practice. It is difficult to explicitly assess the value of information. Often, the value depends on how that information is to be used.

But evaluation planning 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 guideline proposed now lays greater stress on the monitoring of expenditures on evaluation and on keeping within a budget.

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.

The Propriety Guidelines

Discussants often considered that the propriety guidelines presented the greatest challenges of all the four categories of guidelines. The reason for this may be the subjective nature of propriety. What is considered appropriate in one context and culture may be considered serious error in another. Subjective judgments on propriety are determined by cultural values.

Different cultures have different sets of values.

P2.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.

Discussants often 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 and 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 African 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 Kenya Graduate Employment Programme discussants were particularly interested in the possibility of using the guidelines to make the expectations of employers as clear as possible.

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.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. In developing countries, 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 do 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 Guideline on "Rights of human subjects" takes account of these more extensive concepts of rights.

P3.(modified)Rights of Human Subjects.

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.

P4.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.

Discussants noted that cultures exist in which 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 form one culture are interviewing beneficiaries from another. There may be prohibitions against interactions between genders and even between population groups. Some cultural and religious systems place limits on permissible interactions between men and women and for on some topics of discussion also on dialogue between members of the same sex. The type of culturally acceptable interaction is open to interpretation and application.

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.

P6.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.

In several of the review settings there were tensions during the discussion of this PES. Typically, one group would express the viewpoint 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 an advance request to assure the 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.

In most cases, a consensus on the formulation below was achieved in order to maintain hegemony amongst countries while protecting evaluators in Africa required to work with difficult regimes. 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.

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 and without compromising the needs for confidentiality of national or governmental entities.

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.

The notion of conflict of interest is sometimes viewed differently in Africa than in the West. 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 funding?) 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.

The Accuracy Guidelines

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

A1 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 (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 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, though no modification was proposed. Discussants considered that context analysis, like the PES dealing with values identification, is one of that lends itself to international and cross-cultural applications. They noted that many of the subtle nuances of the evaluation would be captured in a context analysis if it were comprehensive. Context analysis would reveal bias and perhaps evidence of situations where certain methods or approaches might not reveal positive results.

A4 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 outsiders. Discussions on sensitive topics such as HIV/AIDS would be strictly taboo. 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 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 information requires a special sensitivity to cultural considerations in some parts of Africa.

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 sensitivities of respondents.

A5 Valid Information.

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

An evaluator who is required to administer a questionnaire on gender specific aspects of behavior to a woman and 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 further 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. 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 and provide answers that would reflect positively on the programme.

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.

A9 Analysis of Qualitative Information.

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

Again, this PES was often discussed though no modification to the Guideline was proposed. Discussants expressed quite a strong consensus opinion that Africa's tradition of passing knowledge down by word of mouth lends itself more readily to qualitative evaluation methods than to quantitative methods.

Conclusions

Formative discussions considered the usefulness of the African Evaluation Guidelines to evaluation work in Africa. In most cases, the African Evaluation Guidelines were considered to be relevant and useful, though

modifications were sometimes suggested. Of the original US PES, 12 were revised and 18 left unchanged. Both political and cultural considerations emerged as major driving forces behind the required modifications.

Guidelines F2 "Political Viability" and P6 "Disclosure of Findings" were both considered politically sensitive in some countries – but not necessarily in all. The guideline formulation used 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. The text of these guidelines was formulated in such a way that it is responsive in both environments.

Cultural considerations were important in the formulation of several guidelines, especially those relating to propriety. P2 "Formal Agreements", P3 "Rights of Human Subjects", and P4 "Human Interactions" all required modification, as did A4 on "Defensible Information Sources". Guideline A5 "Valid Information" was adjusted in consideration of cultural sensitivities to permitted male-female interactions and to queries on topics such as sexual behavior.

In other cases, relatively minor extensions were required. One example is the standard 'U1 – stakeholder identification', which was extended to pay explicit attention to the sometimes ignored beneficiaries at community level.

Many of the US-PES, indeed 18 out of 30, were considered to be fine as they stand and perfectly usable in Africa. The fact that so many were accepted in their current formulation is a reflection of the quality of the initial formulation. And it is perhaps this quality that has made the US-PES attractive to a number of different countries as a basis for developing their own national lists.

None of the existing standards was considered to be totally inapplicable, and no additional standard was proposed for inclusion in the current list. No modification of the category structure of the PES was proposed. Rather, adjustments were made to individual guidelines to make them more readily applicable in current match African cultural, social and political realities.

Two supplementary issues are worthy of consideration here. Firstly, even in this adapted form, the guidelines could still be considered to be a North American derived document. The use of the English language could be considered to result in the embodiment of a number of implicit cultural concepts and assumptions, in some countries. In these, complete Africanisation of the guidelines may require translation into local languages.

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 will be necessary 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 aims to have such a text ready within three years. Until that time, it is hoped that the revised list proposed in this paper will be of some use to those wishing to use a locally developed set of guidelines. It may also facilitate the collection of examples of good practice and the eventual production of a book length treatment of this issue that uses examples from the practice of evaluation in Africa.

Cumulative experience in Africa in detecting areas in which the Guidelines can be improved will serve to further enhance their utility. This adapted list is not intended to be the end-point of an analysis, but rather the starting point of a quest.

Annex 1

The African Evaluation Guidelines 2000:

A checklist to assist evaluators in planning evaluations, negotiating clear contracts and reviewing progress.

The complete working list of the African Evaluation Guidelines is provided below in 4 sections.

Utility:

The utility guidelines are intended to ensure that an evaluation will serve the information needs of intended users.

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 the evaluation findings can be operational 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 contest, 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 e 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 to 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 Subjects. 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.

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 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 and without compromising the needs for confidentiality of national or governmental entities

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 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.

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 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 y 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.

[1] Joint Committee on Standards for Educational Evaluation. (1994). The Program Evaluation Standards. Thousand Oaks, CA. A complete summary of the Standards can be found in Appendix 1.

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

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

[4] Mahesh Patel & Craig Russon. Appropriateness of the Program Evaluation Standards for use in African Cultures. 1998. Presented in modified form to the African Evaluation Association in 1999.

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