

# Utilization-Focused Evaluation in Africa

Evaluation Training lectures delivered to the Inaugural Conference of the African Evaluation Association, 13-17 September, Nairobi

by  
Michael Quinn Patton

Edited by  
Prudence Nkinda Chaiban and Mahesh Patel

## BACKGROUND

The African Evaluation Association was created as an informal network facilitated by UNICEF Eastern and Southern Africa Regional Office (ESARO). The Inaugural Conference of the African Evaluation Association was held in Nairobi on the 13-17 September 1999. It was attended by over 300 evaluators from 35 countries. About 80 papers were presented covering all major areas of evaluation research. Ten national associations or networks of evaluators in Africa were established through this initiative. A database with detailed skill profiles of evaluators working in Africa, created as part of this initiative, is freely available from Mahesh Patel (address below).

The conference had the overall objective of "Increasing Evaluation Capacity in Africa" as well as 6 specific goals: 1) to foster the creation of networks of professional evaluators and professional associations; 2) to develop a sustainable structure to link national associations to an Africa-wide association; 3) to review the US programme evaluation standards for adoption or adaptation in Africa; 4) to create a database of evaluators in Africa; 5) to invite contributions to an annotated bibliography of evaluations in Africa; 6) to publish the papers and proceedings of the Conference.

During the Conference, Michael Quinn-Patton, author of "Utilization Focused Evaluation" delivered a series of excellent lectures. The response to these was so positive that we asked him for permission to make them freely available across Africa as an evaluation textbook. He kindly agreed. The result is perhaps one of the most useful training documents on evaluation freely available in Africa. Please distribute it widely.

The editors would like to take this opportunity to thank the collaborating agencies and donors to the African Evaluation Association. Collaborating agencies were: the African Development Bank, CARE, Catholic Relief Services, Family Health International, United Nations Development Programme, and United Nations Habitat. Financial support was provided by the African Development Bank, DANIDA, International Development Research Center (Canada), Ministry of Foreign Affairs Norway and UNICEF.

Mahesh Patel, UNICEF, Nairobi, December 1999 <mpatel@unicef.org>

Note: Those receiving this document electronically may have difficulties printing the last page of this manual, a very useful flow chart depicting decisions to take when implementing evaluations. To print the flow chart you may need to follow these steps: 1) in the 'File' menu choose 'Print', as usual; 2) in the Print window select the 'Properties' button; 3) in the properties window, select the 'Graphics' tab; 4) in the graphics window click on 'Line Art' and 'Use Raster Graphics'. You may wish to restore your default settings later. The editors apologize for this inconvenience. The rest of the manual should be easy to print. If you have difficulties, please contact mpatel@unicef.org.

Disclaimer: Any views expressed are not necessarily policy of either UNICEF or other sponsoring or collaborating organizations.

## CONTENTS

List of tables 3

Introduction

The beginning of evaluation 4

The broad picture and big pieces of evaluation 5

Standards for evaluation 7

Evaluation in the information age 9

Evaluation as reality testing 10

This refers to the fact that doing evaluations is a way of finding out whether what we

believe is actually true.

Evaluability assessment 13

This is a process that evaluators use, in collaboration with stakeholders, to determine if a program is ready to be evaluated.

Utilization-focused evaluation 16

Key issues and realities in utilization-focused evaluation 16

How to carry out an evaluation process 18

How to identify intended uses and users 19

Process use 23

Leadership training 25

Comments from conference participants 27

Concluding remarks 30

Appendix I 31

Alternative Ways of Focusing Evaluations

Appendix II 35

Examples of situational factors in evaluation that can affect user's participation and use

Appendix III 37

Developing a utilization-focused system for managing outcomes: stages, issues, and activities

Appendix IV 39

Utilization-focused evaluation flow chart

LIST OF TABLES

Table 1.1 7

Ensuring that the big pieces of evaluation are done first

Table 1.2 8

Standards for evaluation

Table 1.3 8

Guiding principles for evaluators

Table 3.1 13

Items on belief in program evaluation from readiness for evaluation questionnaire

Table 4.1 20

Questions to ask of intended users to establish an evaluation's intended influence on forthcoming decisions

Table 4.2 21

Three primary uses of evaluation findings

Table 4.3 22

Examples of situations that pose special challenges to evaluation use and the evaluator's role

Table 4.4 23

Successful and unsuccessful roles of internal evaluators

Table 4.5 24

Examples of the logic and values of evaluation that have impact on and are useful to participants

who experience evaluation processes

Table 4.6 25

Four primary uses of evaluation logic and processes

Table 5.1 26

Four functions of results-oriented, reality-testing leadership

Table 5.2 26

Premises of reinventing government

### **The beginning of Evaluation**

Evaluation was born when someone asked "how do we know what is good?"

Understanding the beginning of evaluation is important because it encapsulates the big pieces of evaluation.

Evaluation is a culture. Anthropologists characterize culture as a group of people who share a language, values, develop norms together, have ways of rewarding positive behavior and punishing the negative, congregate and talk about important things to one another. Evaluators share those characteristics of culture. Other socializing methods of culture are stories, jokes, and metaphors that help us to create mutual understanding. Consequently, most of these have been used to illustrate different evaluation ideas in this paper. The story below, something like the myth of creation, spells out the beginning of evaluation:

The genesis story of Evaluation

In the beginning God created heaven and earth. He looked at everything he had made and said, "this looks very good." On the 7<sup>th</sup> day he rested from all his work and his archangel, or deputy minister in evaluation, came to him and said "God, how do you know that what you created is very good? What is your criteria, and what data do you base your judgement on? What impacts were you hoping to achieve in this creation? Are you not too close to all this work to make an unbiased evaluation?" God thought about these questions all day and his rest was seriously disturbed. On the 8<sup>th</sup> day, God said

"GO TO HELL LUCIFER!" Evaluation was then born.

The story captures the big questions and issues that make up the evaluation field. It is important to make an effort to understand these issues because it is easy to lose sight of them. The big picture may be lost particularly when an evaluator gets down to designing a specific evaluation, at a moment in time, with limited resources.

The genesis story has a fundamental question in evaluation: How do we know what is good? Good is always a value question and not an empirical one. Evaluators can facilitate answers to that question, but cannot answer it themselves. The question on what data are brought to bear on the decision of goodness brings the empirical part of evaluation. Data can be gathered to answer questions on the evaluation process, but values and data have to be brought together to answer questions on whether what happened was good. In the word evaluation is the word value - that is to make a judgement of merit or worth, which requires both values and data about impact, effectiveness, and activities.

In looking at some evaluation reports, the values that inform the interpretations and recommendations are more often disguised than made explicit. And yet part of working in evaluation with people in communities and organizations is to develop standards. It is also to decide what values will guide the evaluation. It is therefore, not enough for evaluators to have skills in research, data collection, and data analysis, but they also must be able to work around people's beliefs and values.

One of the contributions of the United Nations to our global community, as far as values are concerned, is the universal declaration of human rights. From the time the world adopted this declaration, there have been ways of informing the value context of interpreting data that would not otherwise be there. Any training in evaluation that is international in scope should therefore, include issues like universal declaration of human rights and its implications for program and practice. The international value frame provides a way of examining programs especially because a lot of what evaluators encounter involves conflict between different stakeholders (for example international donors and local program leadership).

Another reason why values are so important is that evaluators have learned that, in order for them to negotiate the rough roads that they walk through in doing evaluations, each of them has to have a solid grounding in their own values and ethics that show what is important.

A main purpose of professional associations is to help their membership stay focused on the ethics

and values that guide their practice. That can never be a set of rules. Ethics and values are guided by principles, standards, experiences, and conscience. For instance, even with long experience in evaluation an evaluator is unable to give people rules of how to handle conflicts between donor agencies, local communities, and governments. The evaluator should thus be guided by own values, evaluation standards and principles, and the context, in order to work around conflict and other complex situations.

## **The broad picture and big pieces of Evaluation**

### Some facts about Evaluation

Evaluation is not inherently good. It must deserve its cost

Feeding organizations with reports is not a good reason for the existence of evaluation.

Evaluation builds upon human curiosity and desire to do better.

Evaluation is political because it generates knowledge, and knowledge is power.

Evaluation epitomizes the information and knowledge age.

Evaluation should not primarily be viewed as a technical or methodological enterprise. It is also not about writing reports. Evaluation is really the way people think about what is going on around the world. It is an enterprise that involves thinking. It helps us make sense out of the world - not just related to programs. Out of thinking about what goes on around the world comes the potential of evaluation to empower people. It is therefore, important for individuals and organizations to undertake articulation of evaluation as a way of thinking about the world and a philosophy for engagement in the world, and not as a technical enterprise.

Although evaluation is expanding globally, evaluators understand that the real work and payoff of evaluation is, almost always, at the grassroots levels. Evaluators also understand that there are inherent tensions and conflicts between national and international needs for evaluative kinds of information that inform and empower people at the local level. Managing these tensions of different levels and needs for information is a primary challenge of our times.

The most exciting approaches to evaluation – participatory evaluation and evaluation for empowerment - are emerging from developing countries to North America and not the other way around. These ideas in evaluation are being adopted in North America out of experiences in Africa, Asia, and Latin America. The reason for this is that the more academic approaches that inform the North American evaluation approaches have fallen on hard times there.

One of the present problems in evaluation is that evaluators get caught up in details and forget the bigger and important parts of evaluation. For instance, evaluators usually get started on the evaluation methods (instruments of collecting data) at the first meeting with stakeholders without first examining the purpose, priorities and shared definitions of the evaluation. They usually assume that they know the purpose of the evaluation and that stakeholders understand it in the same way. The down side of this approach is that evaluators, later on in the evaluation process, end up going back to the bigger issues that were not considered at first. Consequently, they often realize that they got off track with the evaluation process because their methods and measurement decisions were heavily determined on what they thought they knew, and not by the information needs of a particular setting.

The following metaphor, which is often used in time management training, emphasizes the importance of doing the big things first:

#### The Jar and Rocks Metaphor

Imagine a large clear jar on your table. You put in it several large rocks until the maximum that the jar can contain. When trainers in time management do that, they then ask the group in the audience whether the jar is full. People typically say "it is full and cannot hold any more." The trainer then takes some small rocks and lets them fall in between the large rocks into the jar. The trainer again asks if the jar is full and the response from the audience is same as the first one. The trainer then adds sand into the jar and it falls between the large and small rocks. Afterwards, the trainer puts in water and demonstrates to the audience that more stuff can actually get into the jar. The trainer asks the audience to think of the importance of this exercise. They say "you can always add more things into your schedule no matter how much you have." They were wrong. The point of this metaphor is that if you do not put in the big rocks first, you can't put them in later. **YOU CAN ALWAYS ADD SMALL STUFF TO YOUR WORK, BUT YOU CAN ONLY DO THE BIG PIECES AT THE BEGINNING.**

The big pieces in evaluation include its philosophy, purpose, and positioning in the world.

Until the big pieces of evaluation are put into place, we do not have any place to connect evaluation to all the other things that go on in and around programs. The foundational pieces, which include philosophy, purpose, and positioning of the evaluation in the world, are required to link evaluation to other things.

In looking at the big pieces, it is important to appreciate that people come to evaluation with different perspectives and backgrounds. Evaluation has therefore, become a rich feast of different approaches. Those approaches depend to a large extent on where an evaluator is positioned, what level of evaluation is being done, and who the primary stakeholders are, among other things.

If you are doing evaluation, particularly in an agency setting, the following suggestions ensure that the big issues in evaluation are considered first.

Table 1.1 Ensuring that the **Big Pieces** of Evaluation are done first

The Do's	The Don'ts
<ul style="list-style-type: none"> <li>&gt; Begin with an interactive session, whenever possible, where stakeholders convene to discuss several aspects of the evaluation.</li> <li>&gt; Discuss the purpose, priorities, shared definitions, who the evaluation is for, and how it will be used.</li> <li>&gt; Match the evaluation to a particular situation.</li> </ul>	<ul style="list-style-type: none"> <li>&gt; Do not discuss the design and measurement instruments in the first meeting.</li> <li>&gt; Do not fall into the rut of doing things just because they have been done before or are easy to do.</li> </ul>

### Standards for Evaluation

What kind of contribution are evaluators making to the human kind, and how will they be evaluated?

In view of the significant expenses incurred by evaluation and the difficulty in carrying them out, evaluators must demonstrate that they are adding value to the lives of people. Otherwise the profession would be cast in doubt. Think about the following example used by UNICEF to demonstrate the high cost of evaluation: Every \$20 used in evaluation is \$20 not available to immunize a child. Evaluation cannot therefore, afford to become wasteful. Evaluators bear the burden of demonstrating that their contributions ultimately increase the number of children who are immunized, hungry people who are fed, productivity of farmers, and more. Correspondingly, standards for evaluation (see table 1.2) refer to what kind of contribution evaluators will make to human kind and how they will be evaluated.

Table 1.2. Standards for Evaluation

<p>&gt; Utility The utility standards are intended that an evaluation will serve the practical information needs of intended users.</p>	<p>&gt; Feasibility These standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.</p>
<p>&gt; Propriety That an evaluation will be conducted legally, ethically, and with regard for the welfare of those involved in the evaluation and those who are affected by the results.</p>	<p>&gt; Accuracy That an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.</p>

It is only human that evaluators have a certain amount of fear at being held to the standards for evaluation. Likewise, people working in programs have fear that an evaluator will come in and judge their worthwhile. Nobody is thus comfortable with being evaluated. For that reason, some guiding principles have been established for evaluators and their audiences to meet the requirements of the evaluation standards. The principles (see table 1.3) also help evaluators to avoid mistakes in applying the standards.

Table 1.3. Guiding Principles for Evaluators

Systematic Inquiry	Respect for people
Evaluators conduct systematic, databased inquiries about what is being evaluated.	Evaluators respect the security, dignity, and self-worth of the respondents, program participants, clients, and other stakeholders with whom they interact.
Competence	Responsibilities for General and Public Welfare
Evaluators provide competent performance to stakeholders.	
Integrity/Honesty	Evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare.
Evaluators ensure the honesty and integrity of the entire evaluation process.	

SOURCE: American Evaluation Association Guiding Principles for Evaluators, Shadish et al. 1995.

In the international sense, it would be useful for every evaluator to pay attention to the universal declaration of human rights. This allows us to uphold our standards in the larger global principles.

One of the advantages evaluators in developing countries might have in trying to introduce an evaluation consciousness, at least in some cases, say at grass root levels, is that people have not already had bad evaluation experiences. Virtually every organization and agency at the communities within North America, which Dr. Patton has worked with, has experienced bad evaluations. In these cases, evaluators have to position themselves as different from their past colleagues, whom the stakeholders might have felt misunderstood them, took away resources, and did not tell their story appropriately.

Some examples of bad evaluations in Africa are the expatriate evaluations where an expatriate evaluator comes in with minimal briefing about the program to be evaluated. The person then spends two weeks doing an evaluation, writes a report, and flies back home.

### **Evaluation in the information and knowledge age**

In the midst of the massive information that we have today, what is worth knowing?

We presently live in the information age, and increasingly the knowledge age. The following is a description of information and knowledge age: A quarter of a century ago people believed that information would be knowledge. They therefore, created huge information systems, which are currently run by computers, to gather data about every possible thing going on in the world. Hundreds of satellites were set up around the globe to gather massive amounts of information about climate, geography, people, and more. Organizations began building massive management information systems to collect data on clients, staff, and anything else that interested them. The challenge, however, was that organizations did not know what information was worth knowing - they had the answers, but not the questions. For instance, organizations had data about every aspect of their clients, but they did not know what they needed to know about them. What is worth knowing? That is the key question of the knowledge age. The importance of this to evaluation is that never in history of human kind, never more than now, has it been true that knowledge is power.

Power has been defined in different ways over the years. According to historians, human beings emerged on earth three-five million years ago in East Africa. During those years humans lived primarily in small hunting and gathering societies. Then came the agricultural revolution, about 10,000 years ago and the industrial revolution, 300 years ago, according to sociologists. We then moved to the information age in the last quarter of century. In the agricultural age, power came with land, while in the industrial age capital was power. In the information age, knowledge is power - people with the right knowledge can get money, land, and everything else that they want. The huge gap of our time is thus becoming the knowledge gap.

In evaluation therefore, as a capacity building activity and not just generation of findings and reports, we address the fundamental questions of our time: What information is worth paying attention to? How do we put information together in ways that make it knowledge? How do we know what is real? What are good and bad data?

It is important to recognize that an evaluator could ask millions of questions at any given evaluation

in the information and knowledge age. Imagine a situation where an evaluator is asked to design a comprehensive evaluation without resource limitations. It means that every question asked would have no methodological constraints. Remarkable designs would emerge in such a situation. These situations are, however, limited or non-existent. The big challenge therefore, is figuring out what is worth knowing and doing that in a way that provides useful information at a moment in time for real decisions. It is not enough to just answer academic questions or find out things that would be nice to know in an evaluation. We need to know what can make a difference in what people do. That is the hard piece of evaluation. It involves preparing people to use evaluation findings and processes.

Think about the fact that a farmer has to prepare the ground before planting seeds. Likewise, evaluators have learned for the past 30 years that the ultimate use of evaluation in the information and knowledge age, whether or not you do something useful, is not determined by the findings that you come up with. It is determined by how the evaluation was prepared from the beginning - this is a statement based on research. This points to the fact that organizations and individuals have to be prepared to use the findings from the beginning of the evaluation plan - which is not natural, automatic or an easy thing to do.

The big mistake that was made in evaluation in the early days was to believe that the findings themselves would carry the day - Power of results! In the knowledge age, however, evaluators have come to understand that organizations are not necessarily thrilled by findings of an evaluation. In fact some organizations have had the following to say when they receive evaluation results:

The report is late and therefore, decisions have already been made;

The report is very thick and nobody will read it;

The questions asked are the wrong ones;

You did not tell us what we needed to know;

The report is full of jargon that makes it boring and hard to understand.

Evaluators thus learned the same lesson that was learned earlier in community development: That work should begin with meeting the beneficiaries; figuring out where they are at; what they are interested in; the ideas they bring to the evaluation process; and then joining professional perspectives to the local perspectives.

#### EVALUATION AS REALITY TESTING

In reality testing, logic and evidence are valued over strength of belief and intensity of emotions. Our minds are, however, physiologically programmed to distort reality as a survival mechanism.

In its most simplistic terms, evaluation is a reality testing enterprise. Reality testing refers to the fact that doing evaluations is a way of finding out whether what we believe is actually true. It is helping people to find out if what they think is going on is really going on. Do human beings need reality testing? There has been enormous breakthrough in brain research on how the brain works and processes information. This research is important to evaluators' utilization of evaluation, which involves how people process information. There are cultural aspects of how we process information that come through our perspectives and socialization, but beyond that there are purely neurological and physiological aspects of how we process information. Apparently, the research shows that in our neurological mechanisms, we are primarily information distorting mechanisms.

There are people called evolutionary neurological anthropologists who are studying the way in which, as we have evolved as a species out of East Africa and populated the world, our minds worked as a result of the selection of the fittest process. Ironically, they are hypothesizing that reality distortion is a tremendously powerful factor in survival.

Research also shows that when people take in any new information, the brain, like a computer, begins a search for where to put that information. It has also been shown that our rational minds are connected to our emotional selves, as we find a place to put new information, in ways that carry the message "I know what is going on here, everything is okay, this feels good."

People who study how the brain works have also developed a number of what they call neurological heuristics. One of them is called the representative heuristic, which means that the brain is programmed to search for the ways that new situations are comfortable, known, and representative of what is already known by the individual. The importance for reality testing in evaluation is that evaluators are involved in what is going on around people. Everybody is evaluating what they do and are deciding how well it is going. People who receive evaluation feedback are prone to sort it out, and figure out where it feels comfortable to tell themselves that things are going well. Cases in point are some organizational cultures that involve people saying, "we know we have problems, but let us talk about the things that are going right." It therefore, takes work and systematic procedures to overcome what appears to be our natural tendency to distort reality and believe what we want to believe.

Learning how to bring people inside reality testing is one of the things evaluators have to offer as a

profession. One of the ways to do that, particularly if working at the community level, is to refuse to let evaluation to be driven by external criteria. Evaluators have to connect with people at the individual level of self-interest - think of how the evaluation will help them. There is a tendency for people at the local level to see evaluation as meant for donors. The local staff are told by their national office, or whatever the mechanism of the organization, that they have to be accountable. They are told that money is being donated and they have to be accountable. As this message gets passed down from level to level, it ends up in some staff or community meeting where evaluation is presented for accountability. How then do you work enthusiasm for evaluation at the local level if you present it in this way?

At every level, evaluation should begin with a positive view of human nature. The best way to present an evaluation at the community level is to tell people that what they are already doing is so important that they owe it to themselves to know if it is working well. It should be recognized that the highest forum for accountability is the individual accountability, but systematic ways are needed to examine that because the individual knows that they should not trust themselves to give an honest answer. A mistake made by many people, however, is viewing evaluation as just report writing, methods, and feeding bureaucracy.

The illustration below is another way that stories and metaphors, that people can relate to in their own lives, help to get people ready for evaluation.

Reality testing story of the emperor's cloth

The emperor called a tailor to get him a new robe. He told the tailor that he wanted the finest cloth that has ever been created. He also wanted it in two weeks. The tailor went off knowing very well that this could not be done in two weeks, but nobody is allowed to tell the emperor that something cannot be done. The tailor went through stress because the emperor had a function where he was going to show off the new robe. After a few days the tailor went back to the emperor and said that the cloth was ready (knowing very well it was not), but claimed that it was so good that it would strike ordinary people dead. The tailor thus preferred to show it to the emperor alone. The emperor sent away all his workers so that he could try the robe. The tailor warned that the cloth was spectacular and could be seen by wise and courageous emperors only. The tailor took out the cloth, but the emperor could not see anything. He, however, could not admit that he could not see the robe else he would be without courage and wisdom. The emperor wore the robe and when he walked down the street, people were asked not to look at him else they would be harmed. People who looked at the emperor as he walked across town passed out, but there was a young child who slipped from his mother, looked at the emperor and said, "THE EMPEROR HAS NO CLOTHES ON."

This story, passed to us from ancient wisdom, challenges us to commit to reality testing, in which logic and evidence are valued over strength of belief and intensity of emotions.

There are many cases where organizations and individuals do not want to hear that things are going wrong in their work (reality). They are interested in good evaluations so that programs can keep receiving donor funds. They turn evaluation into a public relations effort when in fact it should not be a negative thing that a program can fail. We learn more from failure than success. Therefore, mechanisms should be put in place to address failures and avoid repeating mistakes.

There is a need to establish learning organizations, which comprise of people that are not afraid of failure. They view failure as an opportunity for learning. The only real failure is not to learn, and to fail to know what happened, thus repeating the same mistakes without learning from them. Everyone should recognize that the odds against success are high. Goals are not always met and when people believe that they always have to attain success, it leads to all kinds of dishonesty to obtain further funding for programs. Furthermore, in order to get additional funding, people tend to promise more than they can possibly deliver. When the evaluation process comes in, the goals have therefore, not been met because they were not honest to begin with.

## EVALUABILITY ASSESSMENT

It is wrong to assume that all programs are ready for evaluation and that everybody is able to carry out an evaluation.

Evaluators have learned that they would have to carry out the entire program development process in order to achieve an effective evaluation. In fact, distinctions made between needs assessment, planning, implementation, monitoring, and evaluation do not make sense anymore in a world that has become more fluid in reality. The logic of evaluation is what ought to drive all these processes. But since the work involved in all the other stages of program development is not ordinarily given to the evaluator, evaluators have developed euphemism for doing this process over - called evaluability

assessment. It is a process that evaluators use, in collaboration with people working in programs, to decide if the program is ready to be evaluated.

Table 3.1 shows items on a questionnaire on readiness for evaluation - evaluability assessment. These items are as a result of a survey that was developed and is given to program staff to ask how much they agree or disagree with the items. These may not be applicable to all situations, but they are useful guidelines - the instrument went through factor analysis. The more the program staff agree with most items in the questionnaire, the more the program is ready for evaluation, and also the more the likelihood of utilization of results. Thus the questionnaire also helps to prepare people for use.

Table 3.1. Items on Belief in Program Evaluation, from Readiness for Evaluation Questionnaire

Rank order By factor	Item	Factor Loading
1	Program evaluation would pave the way for better programs for our clientele.	.777
2	This would be a good time to begin (or renew or intensify) work on program evaluation.	.732
3	Installing a procedure for program evaluation would enhance the stature of our organization.	.723
4	We don't need to have our program evaluated.	-.689
5	The amount of resistance in the organization to program evaluation should not be a deterrent to pursuing a policy of program evaluation.	.688
6	I have yet to be convinced of the alleged benefits of program evaluation.	-.669
7	Program evaluation would only increase the workload.	-.668

Rank order By factor	Item	Factor Loading
8	"Program evaluation" and "accountability" are just fads that hopefully will die down soon.	-.650
9	Program evaluation would tell me nothing more than I already know.	-.645
10	I would be willing to commit at least 5% of the program budget for evaluation.	.617
11	A formal program evaluation would make it easier to convince administrators of needed changes.	.617
12	We could probably get additional or renewed funding if we carry out a plan for program evaluation.	.587
13	Program evaluation might lead to greater recognition and rewards to those who deserve it.	.548
14	It would be difficult to implement a procedure for program evaluation without seriously disrupting other activities.	-.518
15	No additional time and money can be made available for program evaluation.	-.450
16	Most of the objections one hears about program evaluation are really pretty irrational.	.442
17	Some money could probably be made available to provide training to staff in program evaluation skills.	.385

SOURCE: Smith 1992:53-54

NOTE: Factor analysis is a statistical technique for identifying questionnaire or test items that are highly intercorrelated and therefore may measure the same factor, in this case belief in evaluation. The positive or negative signs on the factor loadings reflect whether questions were worded

positively or negatively; the higher a factor loading, the better the item defines the factor.

Attempting to do evaluation everywhere means that it does not get done well. It is like a sample compared to a census. A sample might be more accurate than a census. Therefore, evaluation is not appropriate for every program and it should not be imposed on programs that are not ready for it. We ought to be opposed to mandated and universal evaluations in programs. It is especially true that most donor agencies make evaluation of programs mandatory. The down side of this approach is that such evaluations produce mediocre results, as shown by research. Furthermore, evaluation is devalued when it is assumed that everybody can do it, particularly when enough resources are not allocated for it.

The alternative to universal and mandatory evaluation is to acknowledge the psychology of human beings: The reaction to being told you must do something is resistance. Resistance may lead to doing evaluations just to meet the requirements of donor agencies. This may lead to dishonesty. Therefore, programs that feel ready to engage in evaluation should submit a proposal to relevant authorities for an evaluation process. The proposal should identify why the programs are ready for an evaluation. Mandatory evaluations end up spending resources to convince those that resist evaluation rather than supporting the early adopters.

If people do not adopt the idea of asking programs to propose their own evaluations, then another suggestion would be to do random samples of evaluations among programs rather than trying to evaluate each and every program. This is to ensure that evaluation is treated as a precious resource, which requires capacity building, and not an extra administrative burden.

It might be appropriate, however, to have minimal monitoring systems in every program to allow basic accountability. In that case, it is important to differentiate between basic monitoring systems and the capacity to do impact and follow up evaluations, which require more support and must not be mandated. The terms monitoring and evaluation can be differentiated, but it all depends on use and meaning given in various settings. For instance, monitoring is known as management information systems in North America - to make it clear that the purpose of those systems is for management rather than for evaluation. Management decisions are made for the smooth running of the program as opposed to evaluation systems that are aimed at fundamental decision-making and program improvement. People have to have clear definitions of what they are doing.

Evaluation is moving from a tools and methods activity to a profession that has a knowledge base of effectiveness. The knowledge base has evolved from the generic effectiveness of programs, in the context of different settings, culture, and needs. Consequently, evaluators, in some parts of the world, are not only being consulted for evaluation, but also in the design of programs. This is especially happening in North America. One justification for this is that evaluators have seen many programs and can help articulate principles of effectiveness and identify, at the design stage, when programs are likely to be going down the path of ineffectiveness. The terms meta-evaluation, meta-analysis, or synthesis evaluation (evaluation of evaluation), which have strong foundations in development work, point to the fact that people look at successful projects in different parts of the world and establish the patterns of effectiveness across those projects, so that other projects might learn and adapt the ideas to their own situations.

In line with transferring aspects of effectiveness from one project to another, there is a whole field of study that has emerged, as part of knowledge age, where researchers study experts to figure out what makes them expert. Part of this research is to develop artificial intelligence - writing computer programs out of the way that experts think. One group that has been studied seriously are the world class chess masters - Researchers know that one of the things that make these people experts is that they can see seven, eight or nine moves ahead in the game.

#### UTILIZATION-FOCUSED EVALUATION

No matter how rigorous the methods of data collection, design, and reporting are in evaluation, if it does not get used it is a bad evaluation. Commit yourselves to evaluations that build capacity.

Every evaluation, in addition to providing quality findings, builds a capacity to think and act evaluatively - and therefore, it gets used. This is important because of evaluating evaluations! The value of an evaluation has to at least equal its cost and should be evaluated according to its utilization. Again the important thing is to build capacity by ensuring use of evaluation process and findings, and not merely providing reports.

#### **Key issues and realities in utilization-focused evaluation**

Commitment to intended use by intended users should be the driving force in an evaluation. At every decision point - whether the decision concerns purpose, focus, design, methods, measurement, analysis or reporting - the evaluator asks intended users: How would that affect your use of this evaluation?

Strategizing about use is ongoing and continuous from the beginning

of the evaluation. Use is not something one becomes interested in at the end of an evaluation. From the moment stakeholders and evaluators begin interacting and conceptualizing the evaluation, decisions are being made that will affect use in major ways.

The personal factor significant to use. The personal factor refers to the research finding that the personal interests and commitments of those involved in an evaluation undergird use. Thus, evaluations should be specifically user oriented - aimed at the interests and information needs of specific, identifiable people, not vague passive audiences.

Careful and thoughtful stakeholder analysis should inform identification of primary intended users, taking into account the varied and multiple interests that surround any program, and therefore, any evaluation. Various stakeholders have interest in evaluation, but the degree and nature of their interests will vary. Political sensitivity and ethical judgments are involved in identifying primary intended users and uses.

Evaluations must be focused in some way; focusing on intended use by intended users is the most useful way. Resource and time constraints will make it impossible for any single evaluation to answer everyone's question or attend to all possible issues. Stakeholders should meet and negotiate what issues deserve priority.

6. Focusing on intended use requires making deliberate and thoughtful choices.

There are various uses of evaluation findings: Judging merit; improving programs; and generating knowledge (see table 4.2). There is also the process use (see table 4.6). Uses can therefore, change and evolve over time as a program matures.

7. Useful evaluations must be designed and adapted situationally. Standardized recipe approaches will not work. The relative value of a particular utilization focus can only be judged in the context of a specific program and the interests of intended users. Situational factors affect use (see appendix II). In conducting a utilization-focused evaluation, the evaluator works with intended users to assess how various factors and conditions may affect potential for use.

8. Intended users' commitment to use can be nurtured and enhanced by actively involving them in making significant decisions about the evaluation. Involvement increases relevance, understanding, and ownership of the evaluation, all of which facilitate informed and appropriate use.

High-quality participation is the goal, not high-quantity participation. The quantity of group interaction time can be inversely related to the quality of the process. In other words, large numbers of participants do not necessarily mean that the range and quality of views will be captured. Evaluators conducting utilization-focused evaluations must be skilled group facilitators.

10. High quality involvement of intended users will result in high quality, useful evaluations. Many researchers worry that methodological rigor may be sacrificed if nonscientists collaborate in making methods decisions. But, decision-makers want data that are useful and accurate. Validity and utility are interdependent. Threats to utility are therefore, threats to validity. Skilled evaluators can help nonscientists understand methodological issues so that they can judge for themselves the trade-offs in choosing among the method alternatives.

11. Evaluators have a rightful stake in an evaluation in that their credibility and integrity are always at risk, thus the mandate for evaluators to be active-reactive-adaptive. Evaluators are active in presenting to intended users their own best judgements about appropriate evaluation focus and methods; they are reactive in listening attentively and respectful to others' concerns, and they are adaptive in finding ways to design evaluations that incorporate diverse interests, including their own, while meeting high standards of professional practice. In this regard, evaluators should be guided by the profession's standards and principles (see table 1.2 and 1.3 respectively).

12. Evaluators committed to enhancing use have a responsibility to train users in evaluation processes and the uses of information. Training

stakeholders in evaluation methods and processes attends to both short-term and long-term evaluation uses. Making decision-makers more sophisticated about evaluation can contribute to greater use of evaluation over time.

13. Use is different from reporting and dissemination. Reporting and disseminating may be means to facilitate use, but they should not be confused with such intended uses as making decisions, improving programs, changing thinking, empowering participants, and generating knowledge.

14. Serious attention to use involves financial and time costs that are far from trivial. The benefits of these costs are manifested in greater use. These costs should be made explicit in evaluation proposals and budgets so that utilization follow-through is not neglected for lack of resources.

### **How to carry out an evaluation process (see appendix III for details)**

**1. Bring together the people that are involved in the evaluation process - the stakeholders.** They represent different constituencies that affect the program. They may include program staff, donors, clients, and others. The stakeholders should then design the evaluation together with the evaluator. This session does not discuss methods of data collection as yet. It begins by examining the evaluation purpose, resources available, and what the stakeholders would like to accomplish in the process. The logic here is very similar to program, project, organization, or community development. The logic shows that before we design interventions (data collection is part of this), we need to understand the context in which the evaluation is taking place. Finally stakeholders articulate their experiences with evaluation and what it means to them. Evaluators should not be banded to the word evaluation - stakeholders might have different words for this process under different settings.

**2. Begin to talk about the profession of evaluation and what it might contribute** - for example provide the standards for evaluation in brief as outlined in table 1.2. Each evaluation associations may have different standards, but the themes are similar. Tell them that the standards are the criteria that they could hold the evaluator accountable for - this is how stakeholders evaluate the evaluator's work. Instead of beginning with their criteria for their program evaluation, begin with the criteria for the evaluation itself.

**3. Explore the intended use of the evaluation by the intended users.** This means that evaluators have to find out the intended uses and users of the evaluation. The evaluation will be evaluated by whether or not it is used by the intended users to meet their practical needs. Evaluation of evaluation is the intended use by intended users. That is the framework for a utilization-focused evaluation. No matter how rigorous the methods of data collection, design, and reporting are in evaluation, if it does not get used it is a bad evaluation. That is the meaning of evaluation standards.

**4. Design instruments and methods of data collection and ask the intended users what they expect the results of the evaluation to be.** The expectations of results should be compared later with the final findings. At the project level, collection of data should be limited to use only - at the grass roots level evaluation is sometimes affected by collecting excessive data in a quick manner. Start small with simple questions to increase learning about evaluation, instead of designing large evaluations that die from their own weight. Posing a single question and collecting data on it excites people about evaluation and answers come quickly this way.

**5. As another test of use, bring the intended users together again and go through some fabricated data.** The intended users will then report what they would do if those were the actual results. This is building the users' capacity to analyze results, helping them to understand data, and increasing their level of commitment to use.

**6. Evaluation findings should be reported in ways that they will get understood and used,** not assuming that they will always take a particular format like a report. They may not even be in the form of writing. They can be oral and interactive. Simplify the presentation of data so that stakeholders can easily interpret it. It is also equally important that the data answers the evaluation question.

Remember that dissemination of findings

is different  
from use of findings.

**7. There should be a follow-up of evaluation after bringing in the findings.** This should normally be included in the initial evaluation budget to ensure that the results continue getting used - it is part of the evaluator's task.

We are trying to illuminate  
what is going on in reality testing  
when we utilize evaluation.

### How to identify Intended Uses and Users

There is a challenge in achieving utilization-focused evaluation. One way of seeking intended use by intended users is to figure out what the decision frameworks are that might be informed by the evaluation. One kind of evaluation use is to inform the decision making process, especially funding decisions. The list on table 4.1 contains questions that an evaluator can ask intended users in order to influence forthcoming decisions. It can also be used to show that a project is meeting its implementation schedule, which is usually an indicator that the project is bound to fulfill its mandate.

Table 4.1. Questions to ask of Intended Users to Establish an Evaluation's Intended Influence on Forthcoming Decisions.

<p>.. What decisions, if any, are the evaluation findings expected to influence?</p> <p>(There may not be any, in which case the evaluation's purpose may be simply to generate knowledge for conceptual use and future enlightenment. If, however, the evaluation is expected to influence decisions, clearly distinguish summative decisions about program funding, continuation, or expansion, from formative decisions about program improvement and on going development)</p>
<p>.. When will decisions be made? By whom? When, then, must the evaluation findings be presented to be timely and influential?</p>
<p>.. What is at stake in the decisions? For whom? What controversies or issues surround the decisions?</p>
<p>.. What's the history and context of the decision-making process?</p>
<p>.. What other factors (values, politics, personalities, promises already made) will affect the decision-making? What might happen to make the decision irrelevant or keep it from being made? In other words, how volatile is the decision-making environment?</p>
<p>.. How much influence do you expect the evaluation to have - realistically?</p>
<p>.. To what extent has the outcome of the decision already been determined?</p>
<p>.. What data and findings are needed to support decision-making?</p>
<p>.. What needs to be done to achieve that level of influence?</p> <p>(Include special attention to which stakeholders to involve for the evaluation to have the expected degree of influence)</p>
<p>.. How will we know afterward if the evaluation was used as intended?</p> <p>(In effect, how can use be measured?)</p>

Utilization -focused evaluation is built upon a lot of practicality, common sense and asking meaningful questions. For instance, we should not pretend that an evaluation done at the end of a project would determine the project's future funding. Real issues must be addressed. An end evaluation, however, may provide lessons learned for application in other programs elsewhere - if that is the purpose, it is a knowledge base use and not a funding one. The utility standard involves asking the following questions: Who are we trying to influence? What are the constraints? What are people willing to change? Evaluators need to differentiate between a program improvement evaluation and a funding decision one, among other types.

Table 4.2 makes distinctions of different uses of evaluation findings. Types of evaluations should depend on the users. For example formative evaluation is most useful when people in the program feel safe to talk about the real problems of the program (needs that are not being met). It works well when there is mutual trust and confidentiality between the evaluator and program staff. A program should not therefore, be jeopardized as a result of stakeholders divulging the problems faced. A way of making formative evaluations work, particularly if the primary users are the program staff, is not to share the findings with donors because they are often misused to punish programs - instead of using them to learn from experiences. But how then do we maintain accountability? **In the context of learning organizations, accountability focuses not on the findings but upon the changes that are made as a result of the findings. This is accountability that is achieved from the use of findings.** It focuses on learning rather than on judgement - this is a paradigm shift.

Table 4.2. Three Primary Uses of Evaluation Findings

Uses	Examples
Judge merit or worth	Summative evaluation Accountability Audits Quality control Cost-benefit decisions Decide a program's future Accreditation/licensing
Improve programs	Formative evaluation Identify strengths and weaknesses Continuous improvement Quality enhancement Being a learning organization Manage more effectively Adapt a model locally
Generate knowledge	Generalizations about effectiveness Extrapolate principles about what works Theory building Synthesize patterns across programs Scholarly publishing Policy making

It is no doubt that evaluators face great challenges when developing the conscience of evaluation use among stakeholders. Table 4.3 elucidates some of these challenges and the special skills that the evaluator may require in addressing different situations.

In addition to looking at different uses of evaluation and discussing the role of evaluation with the users, evaluators need to match the evaluation to a particular situation. There are many kinds of evaluation models for different users. Appendix I lists several types of evaluations that can be done. The challenge is for evaluators to match an evaluation to particular users in a particular situation. Evaluators must therefore, seek to understand the situation in which the evaluation will be applied - that is the expertise to work towards. It is sometimes called situational responsiveness - adapting an evaluation to a situation in the same ways that we adapt programs to situations. There are, however, lots of situations that an evaluator could face. This creates a professional challenge - which is why evaluators are professionals and not merely technicians. They have to think, strategize, have clear sense of direction, and adapt their skills to the challenges. Appendix II provides a list of various situational factors that can affect user's participation and use.

Table 4.3. Examples of Situations that pose Special Challenges to Evaluation Use and the Evaluator's Role

Situation	Challenge	Special Evaluator skills needed
-----------	-----------	---------------------------------

1. Highly controversial issues	Facilitating different points of view	Conflict resolution skills
2. Highly visible program	Dealing with publicity about the program; reporting findings in a media-circus atmosphere	Public presentation skills Graphic skills Media handling skills
3. Highly volatile program environment	Rapid change in context, issues, and focus	Tolerance for ambiguity Rapid responsiveness Flexibility Being a "quick study"
4. Cross-cultural or international	Including different perspectives, values. Being aware of cultural blinders and biases	Cross-culture sensitivity Skills in understanding and incorporating different perspectives
5. Team effort	Preserving credibility	Identifying and using individual skills of team members; team-building skills
6. Evaluation attacked	Preserving credibility	Calm; staying focused on evidence and conclusions
7. Corrupt program	Resolving ethical issues/upholding standards	Integrity Clear ethical sense Honesty

How do we get a sense that the evaluation is going to be useful? This is guided by the analysis of the situation (see different situations in appendix II) and the beneficiaries of the evaluation. The flow charts in Appendix IV provide a step by step utilization-focused evaluation process. One needs to be specific about who exactly is to use the evaluation. The concrete people to use the evaluation must be identified.

To encourage credibility and use of evaluation results, internal evaluators might need to bring in external people, when possible, to provide another perspective on the evaluation process (see table 4.4 for successful and unsuccessful roles of internal evaluators).

Table 4.4 Successful and Unsuccessful Roles of Internal Evaluators

Successful Roles	Unsuccessful Roles
Management consultant	Spy
Decision support	Hatchet carrier
Management information resource	Fear-inspiring dragon
Systems generalist	Number cruncher
Expert troubleshooter	Organizational conscience
Advocate for/champion of evaluation use	Organizational memory
Systematic planner	Public relations officer

SOURCE: Adapted and expanded from Love 1991:9

### **Process use**

This is the most important capacity building development of the last 5 years in evaluation.

Going through the evaluation process can be more useful to people than mere knowledge of the findings. A lot of learning goes on in the process of doing the evaluation itself - this is referred to as

process use. Its conceptualization is new and also fairly controversial. The different steps of the utilization-focused evaluation flowchart in Appendix IV have to do with process use. The process of the evaluation exposes people to think like evaluators - connecting inputs with outputs. Findings and reports come to an end, but what endures in utilization-focused and participatory evaluations is the capacity for people to think evaluatively in program planning, design, needs assessment, and staff job descriptions. Process use is in itself an impact in the form of capacity building.

Table 4.5 shows what it means to think evaluatively.

Table 4.5. Examples of the Logic and Values of Evaluation that have Impact on and are Useful to Participants who Experience Evaluation Processes.

The logic and values of evaluation derive from research methods and communications. These admonitions constitute a "logic" in the sense that they represent a particular mode of reasoning viewed as valid within the culture of evaluation. They are values in the sense that they are what evaluators generally believe. The guidelines and principles below are meant to be illustrative rather than exhaustive of all possibilities.	
Be clear	Be clear about goals and purposes; about what is being evaluated, what data will be collected, what judgements are to be made, how results will be used - indeed, be clear about everything.
Be specific	A favorite evaluation clarifying question: "What exactly do you mean by that?"
Focus and prioritize	You can't do or look at everything. Be intentional and purposeful in deciding what is worth doing and knowing.
Be systematic	Plan your work; work your plan. Carefully document what occurs at every stage of decision making and data collection.
Make assumptions explicit	Determine what can and cannot be subjected to empirical test.
Operationalize program concepts, ideas, and goals	The fundamental evaluation challenge is determining how to measure and observe, what is important. Reality testing becomes real at this point.
Distinguish inputs and processes from outcomes	Confusing processes with outcomes is common.
Have data to provide empirical support for conclusions	This means a commitment to reality testing in which logic and evidence are valued over strength of belief and intensity of emotions.
Separate data-based statements of fact from interpretations and judgements	Interpretations go beyond the data and must be understood as what they are: interpretations. Judgements involve values, determining what is desirable or undesirable.
Make criteria and standards for judgements explicit	The logical mandates to be clear and specific apply to making the criteria and standards explicit.
Limit generalizations and causal explanations to what data support	Overgeneralizations and overly definitive attributions of causality are epidemic outside the culture of research and evaluation.
Distinguish deductive from inductive processes	Both are valued but involve different reasoning sequences.

Thinking and acting evaluatively should be made a part of the program design because it reduces the cost of evaluation and increases utilization of results.

In the evaluation process, people also understand better what their colleagues do and communication among them and beneficiaries is enhanced. This helps people to focus on common goals.

The primary uses of evaluation process are listed on table 4.6.

Table 4.6. Four Primary Uses of Evaluation Logic and Processes

For the uses below, the impact of the evaluation comes from application of evaluation thinking and engaging in evaluation processes (in contrast to impacts that come from using specific findings).

Uses	Examples
Enhancing shared understandings	<p>Specifying intended uses to provide focus and generate shared commitment.</p> <p>Managing staff meetings around explicit outcomes.</p> <p>Sharing criteria for equity/fairness.</p> <p>Giving voice to different perspectives and valuing diverse experiences.</p>
Supporting and reinforcing the program intervention	<p>Building evaluation into program delivery processes.</p> <p>Having participants monitor their own progress.</p> <p>Specifying and monitoring outcomes as integral to working with program participants.</p>
Increasing engagement, self-determination, and ownership	<p>Participatory and collaborative evaluation.</p> <p>Empowerment evaluation.</p> <p>Reflective practice.</p> <p>Self-evaluation.</p>
Program and organizational development	<p>Development evaluation.</p> <p>Action research.</p> <p>Mission-oriented, strategic evaluation.</p> <p>Evaluability assessment.</p> <p>Model specification.</p>

#### LEADERSHIP TRAINING

One way of doing evaluation seminars for senior executives is to call them leadership training - because senior managers do not usually go to evaluation seminars.

Leadership training begins with reconceptualizing how senior management think about monitoring and evaluation by making them see it as an integral part of the program or project. Although called leadership training, it is essentially an evaluation training (see table 5.1).

Table 5.1. Four functions of Results-Oriented, Reality-Testing Leadership

- Create and nurture a results-oriented, reality-testing culture.
- Lead in deciding what outcomes to commit to and hold yourselves accountable for.
- Make measurement of outcomes thoughtful, meaningful, and credible.
- Use the results - and model for others serious use of results.

Focusing on impacts and outcomes is dominant in organizations of all kinds - it is known as the impact mania. It gives an opportunity for lots of bad and good things to happen, and evaluators must differentiate between these. In the leadership training, leaders are taught to see the different types of impact systems. A full picture of the program design is required to know what is going on - relationships between activities and objectives, long and short term outcomes, and impacts. This is the whole point of a logic model.

Establishing impacts largely depends on the degree to which the interventions have a solid research foundation. The greater the interventions are based on solid research of outcomes and effects, the easier it is to see the impacts. For example research has proven that immunization protects children against disease - that is solid research. Research, therefore, provides the linkage between immunization and impact of disease reduction. In areas like community development, the research is

not very solid - there are complex and multi dimensional issues to deal with. In these situations evaluators have to look for important indicators. Evaluators need to differentiate between areas of solid and limited research.

The importance of involving senior people in evaluation is that the unit of analysis for evaluation is gradually shifting from primarily the program/project level to the organizational level. Consequently, program and project success is often hindered by lack of organizational effectiveness. The culture of organizations is therefore, important in the success of programs and projects. The information in table 5.2 is used to train leaders to think about evaluation.

Table 5.2. Premises of Reinventing Government

What gets measured gets done.

If you don't measure results, you can't tell success from failure.

If you can't see success, you're probably rewarding failure.

If you can't see success you can't learn from it.

If you can't recognize failure, you can't correct it.

If you can demonstrate results, you can win public support.

SOURCE: From Osborne and Gaebler (1992: chapter 5, "Results-Oriented Government").

#### COMMENTS FROM CONFERENCE PARTICIPANTS

##### **Comments on the use of mixed methods in evaluation (qualitative and quantitative)**

Qualitative methods make use of open-ended interviews as opposed to survey items that are used in quantitative methods, where questions are closed. Qualitative data is aimed at being rich, descriptive, providing detail, and capturing context, while quantitative data looks at large patterns, generalizes from samples to population, and uses more parsimonious indicators. Quantitative data also captures a lot of information in numbers compared to qualitative data, which provides more depth of understanding in narrative form and is harder to digest, and sometimes to interpret.

A common misunderstanding is that quantitative data are objective and qualitative data are subjective. These terms (objective and subjective) are misleading because neither of the methods is inherently subjective or objective - it is a matter of degree. The interaction between the two methods is important and useful: Quantitative data gives us the large patterns and points to areas where we need more detailed information and understanding; qualitative in-depth thinking then helps us to get additional indicators to understand the big picture and to generate surveys to look at generalizable patterns; we can then do more qualitative research to understand the answers to those surveys.

Therefore, there is a healthy interaction between the two approaches each of which gives us part of a picture, and has weaknesses. The combination gives us a better perspective.

We should, however, desist from believing that the main purpose of integrating these methods is to triangulate. We may get similar findings in both approaches, but triangulation is not a good use of integrating the methods, although this remains a matter of discussion and an area for ongoing learning.

##### **Comments on some old and new management fads**

Each of the following management fads came in to solve problems created by a previous fad:

Early in this century organizations used scientific management to achieve goals and improve themselves. Humans were thus treated and expected to act like machines.

In the 1930s, the human-relations school of thought was born as a result of problems created by scientific management. Organizations began paying attention to their workers and listening to their perspectives and concerns.

In the late 1950s, management by objectives (MBO) emerged and workers had to perform according to objectives. In MBO, however, each worker was working towards their own objective, but nobody bothered to check how the objectives would fit together.

From the 1970s, strategic planning was born as a reaction to the extremes of MBO. Unlike in MBO, objectives in organizations had to fit within vision and mission. Organizations of excellence had to have strategic long-term big picture thinking. Trouble began when these companies, because of their long-term ten-year strategic plans, began to decline in a world that was changing faster than ten years.

The rapid response systems emerged in the information age to complement strategic planning.

Companies that were doing well were moving towards rapid response, rapid changes, constant surveying of the environment, futuring, and were not locked in the ten- year plan. Ten-year plans are not workable in the information age. While the big picture/long term plan was necessary, companies needed to put into account the rapid changes in the world and prepare for rapid response, rapid reconnaissance systems, and on going monitoring systems.

Another notion in the information age is the quality movement and system analysis. Quality and excellence are attributes of successful organizations. This can be achieved through system analysis where leaders have to think and put all the complexity in organizations together. This is where evaluation fits in. Systems thinking is exemplified in the tale of the elephant and the nine blind people. Each of them touched a different part of the elephant and described the elephant based on that part. The moral of the story is that one has to understand all the parts of the elephant and put them altogether in order to understand the elephant. It is a system analytical story - where each part is looked at individually. Analytical thinking involves taking things apart. In order to understand something, we have to look at the larger context within which it operates so that everything has a function within a large system - systems context.

### **Given that evaluation criteria changes, and things in programs change, what is the value of baseline data?**

Part of the way that we know how the world is changing is because we have baseline data. Baselines help us to create trend lines. As data lines are added over time, a trend line emerges - data then becomes more valuable.

### **Comments on evaluation and the project cycle**

It is important to build evaluative thinking in every step of the project cycle, as opposed to having the evaluation at the end of the cycle.

### **Comments on criteria of causality**

Causality is one of the aspects that differentiate evaluation from research. The criterion of causality does not work well under program evaluation because it demands rigorous scientific criteria. The levels of proof required in causality cannot be fully attained in the real world of evaluation. For instance, it is difficult to prove effectiveness in the human world. Instead, we apply the criteria of reasonableness in evaluation - looking at possible causes based on the understanding of data.

### **What is appreciative inquiry?**

Since there can be a lot of negativity in evaluation, appreciation inquiry calls for looking for assets and positive sides of a program. It involves finding out the things that are not working as well as recognizing the assets/positive aspects. It also points to the fact that people's assets can be brought to bear on their problems. The evaluation standards try to achieve a balance in these things.

### **Should evaluators be involved in analyzing the cost side of programs?**

Auditors primarily look at whether funds were used legally and appropriately. They do not technically look at relationships of cost to outputs, effectiveness or benefits. It is therefore, an important methodological contribution for evaluators to look at the relationship between cost and effectiveness and benefits. For instance, start-up costs are always high in relation to outcomes and therefore, cost-benefit ratios low at the beginning of a project. Part of the context of interpreting cost-benefit data is not only to look at the cost-benefit ratio, but also to look at it in the context of the stage at which the project is. It is therefore, an important professional contribution for evaluators to educate people on the different cycles of cost-benefit analysis.

What are the best practices in evaluation?

In the emerging field of systems analysis, researchers have identified how systems operate. One of the principles is that in a well functioning system, no part of the system is operating at its maximum. For example think of the body as a system: If the heart was always operating at a maximum, the body would encounter serious health problems. Also, imagine if the best engineers were convened to hold a competition for identifying automobile parts that are most efficient. If these most efficient parts were put in one car, the car would not run. Each of those most efficient parts is a best practice but only in the context of the car it was designed for.

The best practice of evaluation therefore, has to be applied in context. Evaluations must occur in context and not in routine forms without adapting them to situations in a system.

"Best" comes from old positivist thinking

that for everything  
there must be a best way of doing something.

"Better" is a better word to use.

Comments on the future of participatory evaluation

Bear in mind that in the context of utilization-focused evaluation there is participation of stakeholders. This does not only occur at the community level, but is also applicable to people at every level to think of the evaluation results and how they will be used. Although participatory processes are not easy, and there is more to learn, there are no clear alternatives at the moment for a sustainable development at the community level. There are, however, exceptions to this in the case of targeted short-term projects, which can benefit from a top-down approach: for example immunization.

### **CONCLUDING REMARKS**

The Evaluation-Use Hymn (sang to the tune of Auld Lang Syne)

Written in the point of view of organization executives who sometimes are afraid of what the results of evaluation may show.

May all evaluations done

be used as they should

They tell us how to separate

what is poor from what is good

We gather data near and far

to see what we can learn

The feedback helps us now to know

what to keep and what to burn.

There comes a time for each of us

when doubts can give us pause

We wonder what results will show

will the world see naught but flaws.

But be assured there is naught to fear

if learning is what you seek

Let outcomes guide your every move

Listen to the data speak.

### **APPENDIX I**

Alternative Ways of Focusing Evaluations

Different types of evaluations ask different questions and focus on different purposes. This menu is meant to be illustrative of the many alternatives available. These options by no means exhaust all possibilities. Various options can be and often are used together within the same evaluation, or options can be implemented in sequence over a period of time, for example, doing implementation evaluation before doing outcome evaluation, or formative evaluation before summative evaluation.

Focus of Type of Evaluation	Defining Question or Approach
Accreditation focus	Does the program meet minimum standards for accreditation or licensing?
Causal focus	Use rigorous social science methods to determine the relationship between program (as a treatment) and resulting outcomes.

Cluster evaluation	Synthesize overarching lessons and/or impacts from a number of projects within a common initiative or framework.
Collaborative approach	Evaluators and intended users work together on the evaluation.
Comparative focus	How do two or more programs rank on specific indicators, outcomes, or criteria?
Compliance focus	Are rules and regulations being followed?
Connoisseurship approach	Specialists or experts apply their own criteria and judgement, as with a wine or antiques connoisseur.
Context focus	What is the environment within which the program operates politically, socially, economically, culturally, and scientifically?
Cost-benefit analysis	What is the relationship between program costs and program outcomes (benefits) expressed in dollars?
Cost-effectiveness analysis	What is the relationship between program costs and outcomes (where outcomes are not measured in dollars)?
Criterion-focused evaluation	By what criteria (e.g. quality, cost, client satisfaction) shall the program be evaluated?
Critical issues focus	Critical issues and concerns of primary intended users focus the evaluation
Decisions focus	What information is needed to inform specific future decisions?
Descriptive focus	What happens in the program? (No "why" questions or cause/effect analysis).
Focus of Type of Evaluation	Defining Question or Approach
Developmental evaluation	The evaluator is part of the program design team, working together over the long term for ongoing program development.
Diversity focus	The evaluation gives voice to different perspectives on and illuminates various experiences with the program. No single conclusion or summary judgement is considered appropriate.
Effectiveness focus	To what extent is the program effective in attaining its goals? How can the program be more effective?
Efficiency focus	Can inputs be reduced and still obtain the same level of output or can greater output be obtained with no increase in inputs?
Effort focus	What are the inputs into the program in terms of number of personnel, staff/client ratios, and other descriptors of levels of activity and effort in the program?
Empowerment evaluation	The evaluation is conducted in a way that affirms participants' self-determination and political agenda.
Equity focus	Are participants treated fairly and justly?
Ethnographic focus	What is the program's culture?
Evaluability assessment	Is the program ready for formal evaluation? What is the feasibility of various evaluation approaches and methods?
Extensiveness focus	To what extent is the program able to deal with the total problem? How does the present level of services and impacts compare to the needed level of services and impacts?
External evaluation	The evaluation is conducted by specialists outside the program and independent of it to increase credibility.
Formative evaluation	How can the program be improved?

Goal-free evaluation	What are the actual effects of the program on clients (without regard to what staff say they want to accomplish)? To what extent are real needs being met?
Goals-based focus	To what extent have program goals been attained?
Impact focus	What are the direct and indirect program impacts, not only on participants, but also on larger systems and the community?
Implementation focus	To what extent was the program implemented as designed? What issues surfaced during the implementation that need attention in the future?
Inputs focus	What resources (money, staff, facilities, technology, etc) are available and/or necessary?
Internal evaluation	Program employees conduct the evaluation.
Focus of Type of Evaluation	Defining Question or Approach
Intervention-oriented evaluation	Design the evaluation to support and reinforce the program's desired results.
Judgement focus	Make an overall judgement about the program's merit or worth (see also summative evaluation).
Knowledge focus (or lessons learned)	What can be learned from this program's experiences and results to inform future efforts?
Logical framework	Specify goals, purposes, outputs, and activities, and connecting assumptions; for each, specify indicators and means of verification.
Longitudinal focus	What happens to the program and to participants over time?
Meta-evaluation	Was the evaluation well done? Is it worth using? Did the evaluation meet professional standards and principles?
Mission focus	To what extent is the program or organization achieving its overall mission? How well do outcomes of departments or programs within an agency support the overall mission?