Jenny Hughes

Loek Nieuwenhuis

# A Project Manager's Guide to Evaluation

Evaluate Europe Handbook Series Volume 1

Jenny Hughes and Loek Niewenhuis
A Project Manager's Guide to Evaluation
Evaluate Europe Handbook Series Volume 1
ISSN 1610-0875
2004



Design and layout: Dirk Stieglitz, Bremen, Germany
Printer: Perspektiven-Offset-Druck, Bremen, Germany

This brochure is financed by the European Commission
within the Leonardo da Vinci-Programme.

# Content

# Evaluate Europe Handbook Series

Evaluate Europe is a series of practical handbooks providing resources for evaluators. Evaluate Europe is targeted at those involved in the evaluation process, from professional evaluators to project managers. The series helps to develop a better understanding of evaluation concepts and methodologies and to promote the benefits of evaluation.

What is Evaluation? In general evaluation can be regarded as a joint learning process for all involved, generating useful and relevant information and knowledge. Evaluation is a theoretical and practical approach, which feeds back into ongoing change processes in organisations and projects. It is also a systematic process to assess the relevance, efficiency and effectiveness of policies, projects and programmes. Stakeholders in innovative activities want to know how effective the activities are. What kinds of benefits have been achieved? How can activities be assessed? This all relies heavily on evaluation. The handbook series addresses all of these questions and aims to offer help in formulating new insights for the decision making process in projects, programmes and policies.

Evaluate Europe Handbook Series forms a well grounded resource which will publish work resulting from National and European projects on evaluation. Most of them deal with the design of evaluation processes which are easily adaptable and help to widen the scope of evaluation activity. The series has been initiated by researchers at the ITB, University of Bremen.

The first handbook in the series is "A Project Mangers Guide to Evaluation". This handbook has been written primarily for managers and other professionals responsible for the development and execution of projects funded by the European Commission. (EC) However, much of the content will be relevant to project managers in general and where there are references to EC specific issues, this is made explicit. The key purpose of the handbook is to be a practical 'manual' for managers and an entry-level guide for evaluation practitioners rather than a textbook to be read from the beginning to end. For that reason it has been designed around stand-alone sections so that users can dip into relevant topics.

The handbook has been written by Jenny Hughes (Centre for Research and Education Development, Wales) and Dr. Loek Nieuwenhuis (STOAS, Netherlands). Both have worked together for many years on a range of collaborative education and evaluation research projects and have written extensively on evaluation. They were founder members of CERN (the Capitalisation and Evaluation Research Network). Both have undertaken policy, programme and project evaluations for a wide range of clients including government departments, international agencies, universities and NGOs. They have also been involved in running regular training courses and seminars in evaluation skills for project managers and teaching on post-graduate Evaluation Studies programmes.

Ludger Deitmer, ITB, Bremen, Germany

# Acknowledgements

This handbook has been written with help of the whole CERN team whose ideas and material we have included shamelessly and whose criticisms and feedback have been invaluable.

| | | |
|---|---|---|
| Michaela Woods | IAGO | UK |
| Randolph Preisinger Kleine | P&W | DE |
| Fatima Berger | CESI | FR |
| Bernd Baumgartl | Navreme | A |
| Eduardo Figueira | Academus | PT |
| Nikitas Patiniotis | U.Patras | G |
| Brian Dillon | Nexus | IE |
| Christina Radu | ISE | Ro |
| Helja Fransilla | U.Tampere | Fi |
| Liisa Marttila | U.Tampere | Fi |
| John Konrad | LMU | UK |
| Natassa Kazanzidou | IDEC | GR |

We also want to thank our evaluators Graham Attwell and Tim Sims for their particularly thankless task of having to coax, cajole, plead, bully and finally threaten us into finishing this handbook

We owe an even bigger debt to Dirk Stieglitz of ITB who did the artwork and layout and turned the text generated haphazardly by two computer illiterate people into something readable.

A special thanks to Michael Kelleher with whom we started to write a version of this handbook sometime in the last century and who was happy to let us use anything he had written.

Finally, thank you to all the writers on evaluation whose work we have included, whose web sites we have plundered and who have generously allowed us to use their material. Sometimes we have included ideas and material that we have been using for years and have absorbed into our own thinking, completely losing sight of their source. If anyone recognises their unacknowledged ideas – please let us know and we will gratefully remedy it in the next edition and the web-based version.

There are four sources to whom we owe the greatest debt:
Michael Scrivens at the Claremont Graduate University, California, USA. C.L. Taylor and Larry Arrington and the US Department of Health and Human Services, Administration for Children and Families, USA. Carter McNamara and his work for the Management Assistance Program for Non-profits, St Paul, Minnesota USA. Jody FitzPatrick, (U. Colorado, Denver) James Sanders (West Michigan University) and Blaine Worthen (Utah State University) USA.

We hope this handbook will encourage you to read the original texts!

# Foreword

Evaluation is becoming an increasingly important activity in project management. The emphasis placed on evaluation by policy makers, funding bodies, strategists and practitioners is at an all time high. The cycle of innovating, piloting, evaluating and refining together with dissemination of the process and its outcomes is a widely accepted model of development. Yet many project managers are unclear about what evaluation actually means and, more importantly, how do they do it in practice.

Evaluation is becoming an increasingly important activity. in project management.

A surprising number of management training courses still do not teach the skills of evaluation. Conversely, many standard textbooks on evaluation cover the theoretical aspects of the subject but not the practicalities of evaluation in the context of managing a project. Moreover, the study of evaluation is also comparatively new and there are widely differing theories, models and practices. Evaluation means different things to different people depending on their different needs. We have come across evaluation approaches that contrast starkly with our own but which are nevertheless valid and useful in particular circumstances.

What follows is one approach based on our experience and value systems. We have attempted to bridge the gap between theory and practice and balance the 'how-to-do-it' sections with the 'why-to-do-it' ones. Where we have followed a particular theoretical path or have been selective in our interpretation we have tried to outline the alternative approaches and make explicit the sections where we are being prescriptive with those where we are being descriptive.

This handbook has been written primarily for managers and other professionals responsible for the development and execution of projects.

We hope that this book makes a contribution to both your knowledge and skills as well as providing you with the confidence to make your own decisions based on your own needs and those of your project. In most projects, there is only a restricted budget for evaluation activities. This guide should help you to make informed choices about how to spend that budget for optimal learning output.

Its purpose is to provide managers both with information about evaluation but also to provide practical guidelines on how to do it.

# Section 1

# Chapter 1
# Introduction

## 1.1 What is this guide about, who is it for and what does it do?

Evaluation Mentor software is accessible online and free of charge at: www.evaluate-europe.net

This handbook has been written primarily for managers and other professionals responsible for the development and execution of projects funded by the European Commission (EC). However, much of the content will be relevant to project managers in general and where there are references to EC specific issues, this is made explicit.

Its purpose is to provide managers both with information *about* evaluation but also to provide practical guidelines on *how* to do it or, at least, how to organise it. For this reason, members of evaluation and review committees may also find it useful, as will other project staff who have been asked to take on an internal evaluation role.

Because of the background of the authors and the nature of the Leonardo da Vinci programme, which part funded the production of this handbook, most of the examples are drawn from vocational education and training. However, it is equally relevant to projects in other programmes and contexts.

The first section explores some of the basic ideas and assumptions underpinning evaluation – what it is, why we do it and who the stakeholders are. The second section introduces a simple model of evaluation and looks at the practicalities of conducting an evaluation. The third section identifies some of the common problems and pitfalls and also highlights particular ethical issues impacting on the evaluation process. Finally, we have included a section in which we have collected together some tools, checklists (and further references), which managers may find useful as a starting point for their own evaluation. We have deliberately not included these in the body of the text as they are designed to be photocopied and reused in a variety of situations.

Many project managers will have made the decision to employ an external evaluator. This handbook also includes information on how to employ an external evaluator, how to brief them and what to expect from them

Many project managers will have made the decision to employ an external evaluator – in some cases it may be a condition of funding. However, managers should not assume that this means that they have no further responsibility for the evaluation. This handbook also includes information on how to employ an external evaluator, how to brief them and what to expect from them. We also believe that if managers understand the evaluation process themselves, the more effective and efficient the dialogue between manager and evaluator and the better the evaluation outputs.

The key purpose of the handbook is to be a practical 'manual' for managers and an entry-level guide for evaluation practitioners rather than a textbook to be read from beginning to end. For that reason it has been designed around stand-alone sections so that users can 'dip in' to relevant topics. This inevitably leads to some duplication. It has also been designed to be used in conjunction with the Evaluation Mentor software and follows the same format although both can be used independently.

We would also point out that there are many more sophisticated books and readers on evaluation: for the interested audience we have added a reference list at the end of this guide.

## 1.2 What is evaluation and what it is not

There are probably as many definitions of evaluation are there are books written about it. Here are some of them.

> "Evaluation is the process of making comparisons for the purpose of improving decisions." C.L.Taylor (1999)

> "Evaluation consists of making judgments about programs based on established criteria." Boone (1955)

> "Evaluation uses a systematic method for collecting, analyzing, and using information to answer basic questions about a program - and to ensure that those answers are supported by evidence." ACF Handbook (1997)

> "Evaluation compares what has been accomplished (evidence) with what should have been accomplished (criteria) and then makes a judgment about how well it was done." C.L.Taylor (1998)

> "Evaluation is the systematic collection of information about the activities, characteristics and outcomes of programmes for use by specific people to reduce uncertainties, improve effectiveness and make decisions with regard to what these programmes are doing" Patton (1986)

> "The term evaluation refers to the activity of systematically collecting, analysing and reporting information that can be used to change attitudes, and to improve the operation of a project or programme" Allum (1990)

There are many different definitions of evaluation…all of these are appropriate in different circumstances. It is up to you to work with the one you find most useful.

> "Evaluation does not aim to replace decision makers experience and judgement, but rather offers systematic evidence that informs experience and judgement" Weiss (1999)

> "Synthesizing the definitions from the major dictionaries, we ....take evaluation to be the process of determining merit, worth, or significance. Evaluations are the products from this process." Michael Scrivens (1998)

> "Research determines what can be done, needs assessment determines what should be done, and evaluation determines how well something has been done." Mendenhall (1973)

> "Evaluation is the process of determining the value and effectiveness of a [learning] program. It uses assessment and validation tools to provide data for the evaluation [where]....assessment is the measurement of the practical results of the training in the work environment [and] validation determines if the objectives of the training goal were met." Donald Clarke (1997)

We do not propose to give you another one because we have found all of these are appropriate in different circumstances. It is up to you to work with the one you find most useful for your project. However, the important ideas we can synthesize from all of them are that

– Evaluation is purposeful, it is a means to an end not an end in itself.

– Evaluation of things, which have happened, helps people make decisions about the future.
– Evaluation is based on asking specific questions about a project and finding the answers. It is an investigative process.
– Evaluation is systematic and scientific. It involves collecting evidence, making comparisons, measuring things against criteria.
– Evaluation means that someone, ultimately, has to make judgements about the value or worth of something so its outputs must be interpretive not simply descriptive.

The last definition usefully leads us in to other, related terms, (like 'assessment', 'validation' and so on), that we need to be familiar with in order to understand how and where evaluation fits in.

## 1.3 What evaluation is not.

Understanding what evaluation is not, is a good starting point for understanding what it is!

Understanding what evaluation is not, is a good starting point for understanding what it is! However, because evaluation is a relatively new field it still has an emerging vocabulary and there is considerable variation in the way the terminology is used. Some words are used interchangeably by different agencies, some concepts overlap. For the sake of consistency rather than claiming any authority, we are proposing to use terms in the sense that the European Commission typically uses them – although even that can vary from user to user.

### Monitoring

Evaluation is not the same as monitoring or capitalisation or valorisation or auditing or assessment or...but all of these processes may inform the evaluation process or be informed by it.

Monitoring is about checking - checking whether inputs match outputs, whether income balances expenditure, whether actual activity matches planned activity. It is also about recording the gaps between them. Evaluation is about explaining why the gaps exist. (And remember, the 'gaps' can show over-performance as well as deficits or shortfalls and it is important to record these as the evaluation process will be interested in why this happened and whether there are factors which should influence future planning.)

Monitoring is not the same as evaluation because it is descriptive rather than interpretive and is not intrinsically directed toward learning but the two are often confused. However, evaluation is almost impossible if there is no monitoring system in place.

[Note: In the case of EU funded programmes, internal monitoring systems are the responsibility of the individual project but external monitoring will be usually be done by national agencies through their 'monitoring visits' or through standardised report forms. These are NOT a substitute for evaluation.]

### Capitalisation

Capitalisation has crept into the evaluation-related vocabulary in recent years and simply means building on the achievements of a project (or programme) and using the results in future activities. Thus, project evaluation is a good baseline from which to start the capitalisation process.

### Valorisation

Valorisation is looking at the lessons to be learned from a project and how these can be translated into a wider context or to novel situations. Valorisation is closely related to impact analysis and is usually about the longer-term

sustainability of a policy or strategy rather than a particular application. It is often used interchangeably with 'capitalisation' but is more commonly a term applied to capturing the collective learning of a group of similar projects or a whole programme rather than an individual project.

### Auditing

Auditing simply means 'taking stock of' and is used in this sense in expressions such as 'skills audit' or 'environmental audit' or 'audit of training needs'. However, to a project manager it usually implies an inspection of the project to determine whether financial controls are in place (systems audit) and whether these are being followed (compliance audit). In reality, the terms of reference of the European Commission's auditors or their agents are much broader. They are increasingly concerned with checking out whether there is evidence that the project was needed in the first place, whether the project represents real value for money in terms of its processes and outputs and whether there is 'added value' not only at the level of financial additionality (in a technical sense) but also whether the activity is visible and can be clearly identified in practice.

A companion handbook on 'Preparing for an Audit' is under development in conjunction with Frank McKay, formerly of the Verification and Audit Section responsible for the audit of EU supported projects in the UK.

### Assessment

The difference between assessment and evaluation causes more confusion than almost any other distinction we make. It is largely a question of language and culture and so is particularly problematic on transnational projects. In common usage British English, the terms are virtually interchangeable although 'assessment' tends to be used more in the sense of 'weighing up' or choosing between options rather than in the sense of making scientific judgements about their value or worth. However, in technical usage, particularly in education, training and HRD fields, it is almost always used to refer to the process of measuring the performance of individual staff or students. (This could be, for example, formally, as a result of student examinations or staff appraisal interviews or informally in the workplace.) This is further complicated by the face that in American English, assessment is often called 'Performance evaluation' and recognised as being a distinct sub-category of the overall discipline of evaluation.

The difference between assessment and evaluation causes more confusion than almost any other distinction we make in this field.

So, in the context of a training project, for example, assessment will establish what learning took place whereas evaluation will be asking questions about why that happened.

## 1.4 Why do we evaluate projects?

We believe that evaluation has two main purposes and have called these the 'torch' and 'stick' approaches.

Firstly, evaluation can be about *accountability*. It is a measuring 'stick' that can be used to justify the existence of the project in the first place, its work and its continuation. The ethos is largely inspectorial and judgmental and the underpinning rationale is about value for money, quality standards and is, effectively, a 'licence to practice' for project sponsors. Secondly, evaluation can be about project *improvement*. In this case it can be seen as a developmental process – a 'torch' that helps illuminate problems and recognise good practice. The ethos is diagnostic and interpretive and the underpinning rationale is about collective learning. It is a process that reduces the likelihood of re-

The purpose of evaluation can be about project improvement or project justification.

In the context of a training project, for example, assessment will establish what learning took place whereas evaluation will be asking questions about why that happened.

peating mistakes and using mistakes, when they happen, as critical learning incidents.

Evaluation can be inspectorial or developmental in approach. Both are equally valid, depending on the context and purpose of the evaluation.

Neither purpose is more, or less, valid than the other. The developmental approach is becoming the more popular of the two and many project managers are unhappy with the accountability model. Nevertheless, rigorous 'measuring stick' evaluation is essential if the public interest is to be protected. Not surprisingly, many funding bodies are more concerned with this type of evaluation but are increasingly aware that whilst it is an effective process for justifying existing expenditure, it is less effective for planning long-term investment.

There is also a case for arguing that the dividing line between the two is artificial and that any evaluation should include elements of both. This may or may not be true. However, in our experience, it is not so much that the methodologies are different but more that the spirit in which the evaluation is conducted is different. When we ask clients – usually project managers – the fundamental question "Why do you want to evaluate your project? Is it to justify it or to improve it?" the answer is invariably 'a bit of both'. Whilst that may be true, and theoretically possible, in practice it is very difficult and so our response is usually to say "Fine! But it will cost you twice as much!"

This handbook deals primarily with evaluation for the purposes of project improvement. This is not to say that accountability is not an issue, rather that we have chosen to focus on evaluation as a learning opportunity.

Also, we are aware that our distinction may be too simplistic and that other writers have a more sophisticated break down. Of these, we rather like the Bramley and Newby (in Clarke, D 1995) list, which identifies five main purposes of evaluation – especially number 5!

> "1. Feedback - Linking learning outcomes to objectives and providing a form of quality control.
> 2. Control - Making links from training to organizational activities and to consider cost effectiveness.
> 3. Research - Determining the relationships between learning, training, and the transfer of training to the job.
> 4. Intervention - The results of the evaluation influence the context in which it is occurring.
> 5. Power games - Manipulating evaluative data for organizational politics."

# Chapter 2
# Some fundamentals

## 2.1 Considering the benefits

This handbook is not about 'selling' evaluation as an idea but unless you personally believe that it is a genuinely useful process for you and the project staff, it is unlikely that your evaluation process will be more than another hurdle to jump or simply an administrative procedure. Here are some direct benefits for project managers you might want to consider.

Evaluation should be regarded as more than an administrative hurdle if it is going to support the development of a project.

Evaluation can help you to:

- Find out what is and is not working in your project before too many other people do.
- Show your funding agencies, line managers and the wider community what your project does and how it benefits them.
- Raise additional money for your project by providing evidence of its effectiveness
- Improve your staff's work with participants by identifying weaknesses and strengths thus contributing to the staff development process.
- Improve your personal credibility and reputation by adding to the existing knowledge in the field in which you are working (e.g. in terms of what does and does not work in your type of project with your kinds of participants as well as in terms of outcomes).

## 2.2 What are the basic questions an evaluation can answer?

There are many different types of project evaluation, many different terms to describe them and many questions that they can answer. You may have heard the terms formative evaluation, summative evaluation, process evaluation, outcome evaluation, cost-effectiveness evaluation, cost-benefit evaluation and impact analysis. Definitions of these terms and others and selected resources for more information on various types of project evaluations are provided in the toolbox.

Different types of project evaluation and their associated terms are summarised in the Toolbox and in the Glossary.

You may have also heard the terms 'qualitative' and 'quantitative' used to describe an evaluation. However, these terms, which are also defined in the glossary, refer to the types of information or data that are collected during the evaluation and not to the type of evaluation itself. For example, an outcome evaluation may involve collecting both quantitative and qualitative information about participant outcomes.

This handbook is designed to avoid the confusion that often results from the use of so many terms to describe an evaluation. Instead, all of the terms used here are directly related to answering evaluation questions derived from a project's objectives. There are two types of *project objectives*:

### Project implementation objectives

Project implementation objectives refer to all those things that you plan to do in your project. They will cover how you plan to do it, who you want to reach, the types of activities you will include, the services, training you want to provide or the research and development work you want to undertake. They should also refer to the characteristics of the participant population, the

SMART objectives
Specific
Measurable
Achievable
Realistic/relevant
Time bound

number of people you are targeting, and the time scale. You may also need to specify other processes at the planning stage e.g. the staffing and staff training arrangements, strategies for recruiting participants, number and type of meetings, management and co-ordination systems publicity and so on. This information is needed for budget calculations and, therefore, has to be built in at the outset. Evaluating project implementation objectives is often referred to as process evaluation. However, because there are many types of process evaluations, this handbook will use the term *implementation evaluation*.

Participant outcome objectives.

Participant outcome objectives describe what you expect to happen to your participants or beneficiaries or target groups as a result of your project, with the term 'participants' referring to agencies, communities, networks and organisations as well as individuals. Your expectations about how your project will change participants' knowledge, skills, attitudes or awareness are your participant outcome objectives. Evaluating a project's success in attaining its expectations for participants is often called an *outcome evaluation*.

An evaluation can be used to determine whether you have been successful in attaining both types of objectives, by answering the following sorts of questions:

– Has the project been successful in attaining the implementation objectives? (Are you implementing the services or training that you initially planned to implement? Are you reaching the intended target population? Are you reaching the intended number of participants? Are you developing the planned collaborative relationships? Etc.)
– Has the project been successful in attaining the predicted participant outcome objectives? (Are end users exhibiting the expected changes in knowledge, skills, attitudes, behaviours or awareness?)

A comprehensive evaluation should look at both the success of the project in meeting its implementation objectives and participant outcome objectives.

A comprehensive evaluation must answer both key questions. You may be successful in attaining your implementation objectives, but if you do not have information about participant outcomes, you will not know whether your project is worthwhile. Similarly, you may be successful in changing participants' knowledge, skills and attitudes but if you do not have information about your project's implementation, you will be unable to identify the parts of your project that contributed to these changes.

These are all the sorts of questions that project managers and staff ask and answer on a routine basis, totally informally. Are participants benefiting from the project? Are the strategies for recruitment working? Are participants satisfied with the services or training? Do staff have the necessary skills to provide the services? Are we up to speed on the development? Is the management team working well? How is the research progressing?

Evaluation addresses these same questions, but *uses a systematic method for collecting, analyzing, and using information to answer basic questions about a project* – and to ensure that those answers are supported by evidence. This does not mean that conducting an evaluation requires no technical knowledge or experience - but neither does it mean that evaluation is beyond the understanding of project managers and staff.

## 2.3 When do you start?

One goal of evaluation is to fix problems and make the system better, not to lay blame – so these evaluation questions should be answered as far as possi-

ble while a project is in operation, not after the project is over. This will allow you and your staff to identify problems and make necessary changes while the project is still operational. It will also ensure that project participants and partners are available to provide information for the evaluation. This is often called *formative evaluation* as opposed to *summative evaluation* – which is the final-review type of evaluation undertaken at the end of a project.

An Effective evaluation should begin as early in the project as possible and not after it has finished.

(Personally, we do not find this distinction helpful if the evaluation is about project development rather than project justification. If the evaluation is essentially about accountability, there is a case for a summative report that 'signs-off' the project. However, if the project is about learning, then even a final report should be constructed in a formative way around future improvements – a baseline for new work.)

Ideally, the evaluation process should start when the project is still on the drawing board. Designing the evaluation is an integral component of the overall project design and should be built in to the planning stages. In fact, most funding application forms ask for a description of the proposed evaluation process. However, in reality this has one major drawback. Because of the long lead-in times, the group of people who draw up the funding application may not include the project manager and neither the project manager nor the project designers may have experience of evaluation design. If this is the problem, the solution is often to appoint an external evaluator. However, external evaluators cost money and usually there is no money to appoint them until after the application has been approved. This means that their expertise in drawing up evaluation plans is already too late!

Technical expertise may be needed in the early design and operational phases of the project.

Under the section on 'Using an External Evaluator', we have suggested some ways around this. You may also like to use the 'Mentor' software which is a series of structured questions designed to help you make decisions about evaluation as early as possible. Part two of this handbook also provides some practical help to get you through the early design stages.

Once the project is approved and underway, the evaluation process needs to run in parallel with its implementation from the first day. Ideally it should not be a 'bolt-on' to the other activities but integrated within the mainstream project processes. Again, ideally, the evaluation should continue after the main activities of the project have been completed in order that an *'impact analysis'* can be carried out - that is, a view of the longer-term consequences of the project and its sustainability rather than its immediate outcomes. In practice, this is hampered by the time-bound conditions of funding but may be possible if there is a second phase or follow up project, which could build in resources for carrying out this type of evaluation of earlier work.

The evaluation process should be integral to the project's day-to-day operation - from the first day.

You may come across the following terminology relating to evaluation carried out at different stages in the project process. Although these are terms that suggest discrete evaluation activities, it is probably more helpful to stick with the idea of evaluation as a continuous and on-going process and to consider these simply as reporting phases. The only time when they are likely to be organised as completely separate events is in the evaluation of large-scale programmes when there are additional issues around accountability. As a project manager, these have no direct relevance, other than a familiarity with the terms is useful.

### Ex ante evaluation

This is the initial scoping study or position audit, which describes the project or programme's starting environment and provides the baseline information for future evaluation. It enables the project evaluators to measure the 'distance travelled' (see also 'ipsitive referencing').

### Mid term evaluation

This is the point at which progress-to-date is formally measured to see whether the original environment has changed in a way which impacts on the relevance of the original objectives. It is an opportunity to review these objectives if necessary, decide whether the project is on target in terms of its projected outputs, adjust the working practices if necessary or, in certain circumstances, re-negotiate timescales or outputs. It is often not carried out at the 'midpoint' at all but at the end of a significant phase!

### Ex post evaluation

This is the evaluation produced after the project is completed, which includes not only the summative evaluation of the project itself (typically in terms of processes and outputs) but also an analysis of the project's impact on its environment and its contribution to wider (economic/societal/educational/community etc.) goals and policies. It should also lay down a framework for future action leading, in turn, to the next ex ante study. In reality, ex post evaluations often take so long to produce (in order to measure long-term impact) that they are too late to influence future planning.

# Chapter 3
# Organising an evaluation

## 3.1 Getting started - coping with a project application form.

Develop a detailed evaluation plan for your project as part of work package number one.

As we pointed out in the last chapter, evaluation ideally needs to start during the design stages of the project. However, the good news is that you do not need a detailed plan at this stage. An outline plan is usually sufficient for most project applications. What is important is that you include as part of your first work package the task of developing the evaluation plan in detail - and meaning it! If you consider the following questions, the answers will give you an outline plan which meets the requirements of most funding bodies at the application stage.

## 3.2 What is the purpose of the evaluation?

Who are the recipients of the evaluation products? The answer will determine the scope and method of the evaluation process.

An opening statement on the prime purpose of the evaluation is essential – is it directed towards accountability issues, is it to be primarily a developmental process or are you intending to do both? (In which case, reflect this in the evaluation budget). To help you answer this question, it is useful to think who the evaluation is for, that is, who will be the prime recipients of the evaluation products and to explain this in the application. (Is it, for example, for the funding agency or the promoters or for policy makers or the management team or for the participants or partners?) By and large, practitioners will be more interested in evaluation as a diagnostic or learning process, whereas project sponsors are more likely to be interested in a project justification-approach. (More help on this is given in chapter 5)

## 3.3 What are the questions that you want answered by an evaluation?

Try writing a statement in the application form that starts *"By the end of the project the evaluation will answer the following questions..."*. Then you can list them. These questions should reflect the purposes of the evaluation you have already specified and reflect the type of project and its objectives. For example: What impact has the training had on the organisation? Was a return on our investment realised? Has the project changed attitudes, behaviours or skills in a way that positively impacts on business results? Has the research influenced practice? Are the learners using their new techniques and processes back in the work environment? Is the network sustainable? These are examples - forget just filling in the application form, you really need to know what these questions are at the project design stage because it will affect the costs of the project and its core activities.

## 3.4 How comprehensive will the evaluation be and what is the general approach?

You cannot evaluate every aspect of the project as the time and resources would be prohibitive. You need to provide in the application form some idea of the scale and centrality of the evaluation process. Is it to be a major investment, contributing significant added-value to the learning process and a major role for the evaluator(s)? If so, you need to justify this. Is it a process that

At application stage, decide what aspects of your project need evaluation. Balance the focus and scale of evaluation with resources and staffing capacity.

will be embedded in other activities of the project (for example, as part of piloting services or products)? Explain this and be careful you do not 'double-count' the costs. Will it be a process of self-evaluation and internal review? This may require extra meetings, which will have to be justified and you will need to convince the funding body that you have the skills and experience to do this.

### 3.5 What will be measured?

Are you intending to look only at project outputs (see *participant evaluation*) or at the internal working processes of the project (*implementation evaluation*) or both? If you intend to evaluate the processes, there is a list of possibilities in section 6.3 (under '*Dimensions of Performance*') but this is ONLY a list of suggestions. Don't assume you have to do all of them – better to do fewer well than claim you will do everything.

### 3.6 Who has authority and responsibility for the evaluation?

As part of the evaluation design, clearly define organisational roles and responsibilities, that is, who will organise and who will undertake the evaluation .

State whether this is the project manager, director or co-ordinator, an evaluation committee, someone else in the promoter organisation (e.g. a quality control department) one of the partners, an external evaluator, an internal evaluator or a combination of these. (In which case, explain the relationship between them). It is important to explain who will have responsibility for actually organising and undertaking the evaluation in practice and who has the responsibility for commissioning them and providing their terms of reference – that is, to whom are they accountable. You will also need to say whether they are sub-contracted or included in salaried staff. If you are not using an external evaluator on a big project, you will have to justify this – maybe because of the evaluation methodology (e.g. self evaluation, but see above) or because of 'independent' in-house expertise (as in a specialist department in the organisation or because of the nature of the project (e.g. software projects might build in evaluation as part of an iterative development and piloting process.)

If one of the project partners is to take on the responsibility for evaluation, this should be explained in the work packages AND in the evaluation section of the application form and will need a brief explanation of how the integrity of that role is to be protected (e.g. they should not be involved in other project activities).

### 3.7 What will be the role of evaluator

Detail in the application form any proposed evaluation activity that may be interpreted as 'additional' to general evaluation activity.

This does not need explaining unless the role is significantly different or 'enhanced' to include other related activities. For example, some developmental evaluation processes could ask the evaluator to take on a project team 'mentoring' role or set up a 'rapid feedback' system – almost a consultancy facility for project managers. Evaluators could be facilitators at team review meetings or act as a 'critical friend' to project staff or take on the role of feeding in ideas or research from other projects. All of these are legitimate evaluation activities but will need justifying – particularly as they require additional skills and expertise on the part of the evaluator.

### 3.8 What will be the main sources of the data and how will the data be collected and compiled?

This does not need to be a detailed methodology but should give an idea of whether, for example, large-scale postal surveys are involved, extensive interviews, focus group meetings or the use of statistical analysis. All these will have cost implications and it might be necessary to explain this to justify expenditure on meetings, communication costs, statistical software packages and so on, over and above the evaluators' fees

### 3.9 How will the data be analysed and presented?

Make it clear in your application how the evaluation data will be collected and provide a timetable for the distribution of evaluation products.

The application form should explain who will be involved in the data analysis and preferably what arrangements are in place governing the authorship and ownership of the evaluation products. This will give some indication of the authenticity, reliability, validity and objectivity of the evaluation outputs. You will also need to specify the type and number of evaluation products and when in the time scale of the project they are to be produced e.g. reports, seminars etc. Again these should be reflected in the budget. (For example, if there are to be 100 copies of a final 'glossy' evaluation report, the responsibility for production, reprographics and distribution could be included in an external evaluator's sub-contract or the evaluator could just be required to produce a disk copy and the project co-ordinator does the rest. This will need explaining to justify budget lines.)

### 3.10 How much will it cost? Have you enough resources?

There is no need to include the costs of evaluation in the main body of the application or in the section of it relating to evaluation, as it will appear in the budget pages. However, you should explain anything out of the ordinary (e.g. very high or very low expenditure on evaluation). It is almost impossible to say what a 'reasonable' figure is as circumstances vary so much across projects. However, as a very rough rule-of-thumb, expenditure on evaluation for projects up to 100,000 € should be about 8-10% of the total budget. Between 100,000 € and 250,000 € about 5-8%, between 250,000 € and 1 million € about 5% and over 1 million € about 3%. Obviously there will be anomalies at either end of these scales but these figures are not a substitute for accurate costing, only ball-park figures to give you a 'feel' for what is 'high' and 'low'. One thing to remember when allocating resources to evaluation are the hidden costs such as internal staff time needed for interviews, data collection, additional travel, reprographics and so on.

Although a breakdown of evaluation costs may not be needed, base your estimates on a solid rationale.

You are not usually asked to break down the costs of the evaluation for the purposes of a project application – for the most part they can all be included as a single budget line, (especially if it is a sub-contract) unless there are evaluation costs which impact on other budget lines (see some of the examples above), in which case this probably needs to be explained. However, YOU need to be able to calculate the costs and to know what is realistic.

Finally, remember that all of the questions above are real issues which have to be considered as part of the project design and costing at the application stage and not simply a formula for completing the relevant section of the form – although they may provide you with a checklist of what to include and a way of laying it out. In addition, it will give you the basis of a specification for your evaluation requirements if you are proposing to tender to an outside evaluator. (see also the section on 'Working with an outside Evalua-

tor') Usually your evaluator will refine your outline plan and add a detailed methodology and this needs to be done as soon as possible after the project is approved. If there are major changes to the outline plan, then you need to notify or negotiate this with your funding agency.

# Chapter 4
# Evaluating European sponsored projects

This chapter sets out some of the important characteristics of EU sponsored projects and their implications for evaluation. If you are not involved in managing or evaluating one of these projects, then you can ignore this section and move directly on to chapter 5.

We have identified 11 key features of EU sponsored projects and considered how these impact on the evaluation process.

### Most, if not all, programmes, ask for evaluation plan to be built into application forms together with performance indicators and a compulsory budget line for evaluation (1)

This is good news and reinforces our own beliefs that evaluation should be properly funded and built in at the front end of projects, not bolted on at the end. However, this is not without its own problems.

As projects are often managed by people who have little or nothing to do with the application, evaluators also frequently work without a sense of the history of the project. Where has it come from? What experiences have led this organisation to seek support for such a project? Etc. Are you evaluating against the project as conceived at the time of application or at approval or as the project ought to be now?

Firstly, the project application is usually written by someone in the promoting organisation who may or may not have anything to do with the management or evaluation of the implementation of the project and may know nothing about evaluation as a process. (The section on 'Getting Started' is designed to help) Appointment of a project manager may actually depend on the project being approved, so they will take up their position with some major decisions about evaluation having already been made.

Similarly, the evaluator, especially if this is to be an external evaluator appointed on the basis of a tendering process, cannot inform the process from the outset. This can be limiting for the evaluator once appointed as they may have less influence or control over the process for which they are responsible. It can also create unwilling clients who just see this as another hoop to jump through.

In ideal world, evaluators should be involved during the preparation of the application. You may find that an evaluator will do this for no fee, as long as they know that they will be asked to do the job if the project is approved, which is not unreasonable! This at least ensures that the planning and implementation of the evaluation is coherent. However, the European Commission takes the view that all project sub-contracts should go out to tender if they are 'substantial' (Figures not usually specified, but at the time of going to press, assume that any external evaluation contract in excess of 20,000 € will cause raised eyebrows if there is no tender process. Your own organisation (particularly the public sector, may also have procurement regulations).

This is a difficult situation and one that the technical assistance units should consider. However, there are some quick-fix solutions.

– Nominate an evaluator at the application stage, enclose their CV if the application process asks for details of key staff) and justify this (if its true) by saying that they have been already been included at the project preparation stage to advise on the evaluation process from the outset. Demonstrate that this is good practice because evaluation has been built in from the beginning

– Make one of the partners responsible for evaluation or, to put it another way, make your evaluator a partner. You will need to say how you will protect the integrity and independence of that partner, what the reporting mechanisms will be and what the relationships will be between that partner and the co-ordinating partner. The evaluator-partner should not be involved in any other project activity – except, possibly, informal dissemination.
– Pay an evaluator to work with you at the design stage, complete the evaluation section of the project application and prepare the tender specification for external evaluation on the understanding that they would not be eligible to tender. This is about 2 days work.

Secondly, applications forms frequently reinforce a model of evaluation that is not integrated with the rest of the project activities. Diagrams of project management structure, work plans and deliverables rarely show the relationship with the evaluation aspects of the project. Evaluation plans are sometimes confused with monitoring sections of the forms and exist separately from other important sections on the project management and methods.

### European projects operate within programmes that generally have a fairly sophisticated quality assurance and accountability infrastructure (2)

There are many agencies and structures that have roles which potentially impact on the evaluation process. National Co-ordinating Units, Technical Assistance Units, Verifiers, Auditors, Monitoring Committees etc. are all involved with activities that are concerned with ensuring the quality and value for money of projects. (The differences between these processes are outlined in section 5.1) This has consequences for the ethos, relationships, objectives and processes of the evaluation. It is important for project managers to understand how these roles are differentiated what the relationship is between them.

There is a potential danger of treading on each other's toes. The individual agencies may be clear about what roles they play, but is it clear to managers of the projects? The potential for duplication is evident yet it may also be possible to leave gaps. What access does the project evaluator have to the overall work of these agencies? There are no clear answers as the roles of the different agencies will vary from programme to programme. However, as a generalisation, you can assume that they will be primarily concerned with ensuring projected outputs are delivered, there is financial accountability and a justification for the project's existence. You may want your evaluator to check on these too but the essential difference is that the added value of the evaluation process is that it is about learning.

### European projects are time bound and output based and these must be stipulated at the outset (3)

For many reasons, the project illustrated in your application is rarely the project you are able to deliver. An effective evaluation will identify areas of divergence and will allow the project manager time to notify the relevant funding agency of any potential significant changes.

All projects will start off with a clear plan and precise objectives, which will form the basis of the evaluation process. In most cases the project funding will be dependent on achieving specified outcomes. However, the world changes. An evaluation process that is built in from the start of the project may help identify situations in which the original objectives are no longer relevant. There may be changes in the project environment, for example the labour market may have changed significantly or the pool of beneficiaries may have changed making it difficult to recruit or retain which in turn means the out-

comes and deliverables may need adjusting. Or there may have been changes in the operating environment where the feedback from staff, beneficiaries and the evaluation team suggests that there are different and more effective ways of doing things or priorities need shifting because different needs are identified as the project has developed.

Ironically, the more effective and more rigorous the evaluation processes – in other words, the more critically reflective the project – the more the plans are likely to need changing. However, project managers often assume there is an in-built inflexibility in the administrative systems that make changes impossible. Our experience is that funding agencies are very open to changes being made IF it can be clearly demonstrated that external circumstances have changed or that the project is addressing emerging issues. Funding agencies want projects to succeed. For the most part they accept that projects should be dynamic, flexible, responsive and continually improving so that there is a best-fit between project deliverables and the needs of the communities the project is serving. It is often the case that the original objectives can stay the same but the outputs and methods may need to be renegotiated. The key is to make sure that relevant agencies are involved as soon as possible and that they are part of the discussions rather than simply being notified - or worse, not told at all. A properly conducted, preferably external, mid-term evaluation' report which sets out what changes need to be made and why can be incredibly useful in persuading other agencies that changes to the project should be approved and is evidence that the proposed amendments have been properly researched.

Many EU programmes are not good at dealing with 'failure' – in other words, projects cannot fail without incurring financial penalties. However, It can cost as much to 'fail' as to succeed if the project is genuinely putting maximum effort and resources into trying to make the project work. So another effect of output related funding is that promoters are often tempted to go for the lowest acceptable level of output in order to minimise failure. The by-product of this system is that is inhibits ambitions and places a glass ceiling on projects. We end up with less challenging, less innovative outputs.

A parallel problem is that most programmes demand demonstrable and quantifiable outputs in terms of products or people. Very few allow for legitimate 'process' outputs. Thus, there are no formal mechanisms for learning from processes and much of the process-based learning involved in evaluation may be lost to the project and to a wider audience.

### EU funded programmes operate in a political context which has implications for projects and, therefore, for their evaluation (4)

Projects exist within the context of programmes, which in turn exist in a changing EU policy environment. Each programme will have a general set of policy objectives and mechanisms for translating these into practice through the projects it supports. However, the policy directions are constantly shifting. For example, the original Education, Training and Youth programmes supported projects that reflected member state vocational education policy. Increasingly, there has been shift towards supporting projects that promote European policy rather than national policies. This has implications for the evaluation questions about a project's effectiveness in meeting its objectives, certainly from the perspective of the paymasters – who are major stakeholders.

In addition, VET projects may be connected to lots of other policy fields such as regional development, labour market development, social inclusion and so on, which should be reflected in the evaluation.

### EU projects must demonstrate in some way, the added value of the project at a regional level (e.g. for the structural funds) or a European level (e.g. for the Education, Training and Youth Funds) or at occupational or sectoral levels (5)

This is a major difference as many non-EU funded projects only want an evaluation to show how effective the project has been within its own organisation or immediate environment. EU project evaluation, however, needs to be conducted locally but at the same time think of the consequences and potential applications further up the value chain. Extracting transferable issues may be the role of the evaluator

### Many European funding programmes have a heavy emphasis on innovation (6)

For further information see 'Evaluating Innovation –describing the process', Hughes and Attwell 2000, www.cred-wales.co.uk

EU funding is often pump-priming money to develop new approaches and novel products. This begs the question, "Should 'innovativeness' be a specific element of the programme to be included in the evaluation or should the evaluator be satisfied that evidence of the innovative nature of the programme was provided at the project application stage and accepted at the project approval stage?". The answer probably lies in who the evaluation is for. Either way, this is an issue which must be discussed by the project manager and the evaluator from the outset.

### In many EU projects there will be a compulsory transnationality element (7)

How do evaluators handle this? Is transnationality an aspect of the project that must be included in the evaluation or is it optional? Is it the whole project or the national aspects of it that is subject to the contract for evaluation? What are the key evaluation questions around transnationality? How central is it to the project? What is the model of transnational working and what are the processes? (collaborative research? a network? joint development of products and services?)

Where funding for a transnational project is at the national state level, (for example, as it was for the Community Initiatives), who pays for transnational evaluation? With whom is the contract? Who is the client?

### Multi-annual programmes often depend on outcomes at each stage in order to call down payment and a unit-costing model does not necessarily reflect the project life cycle (8)

This is particularly true of ESF type projects where there are often tensions between building solid foundations and a funding model that frequently demands outputs in the first phase of the project. The most intense activities of the evaluator, who will be central to making sure that the project is on course, will be as early as possible, especially if the evaluator is involved in carrying out an ex-ante evaluation. Tensions might exist in the building of foundations

### The applications forms are helpful in that they are not prescriptive about evaluation. You must have it but what, how, when and by whom is a matter for you (9)

From a project perspective, this can be a good thing as it provides you with autonomy and flexibility and opportunities to explore different evaluation models and work with different evaluators. However, from a programme perspective it makes aggregating evaluation outcomes from the different projects extremely difficult.

However, the Leonardo da Vinci E-VAL2 project has made inroads into solving this problem with an electronic tool that codes text across evaluation reports. Contact the authors for more information.

A related issue is that evaluators will be working with possibly more than one project and common issues do arise in evaluating projects. Information and knowledge stemming from the wider project community may have some bearing on the evaluation processes and judgements of a particular project. This may be useful or it could result in role contamination as well creating legal and ethical issues that prevent effective evaluation to be carried out. (see also Chapter 11 on ethics)

Programme valorisation is being developed to address the collective learning of EU projects but it does not use the collective knowledge held by project evaluators.

It also follows that evaluators could be expected to support the collective learning of all projects by making those generic lessons learned available to policy makers and to share their knowledge across projects without disclosing confidential information. (see also Chapter 11 on confidentiality.) To date no mechanisms exist to support that sharing. There may be a good case for each programme to create a network of project evaluators to support this work on behalf of the projects

### European projects have a lot of stakeholders (10)

We explore in section 5.1 the issue of relationships between stakeholders in the evaluation process. The contract for evaluation will typically be with the project sponsoring body but it could be part, or even, wholly, funded by the European Union and they may be the agency asking for the evaluation to be carried out in the first place. They may specify that it must be done and require you to say how it will be done, but do they have any further interest in the evaluation outcomes? In other words, who are the clients?

Our own approach to this is relatively straightforward. If the evaluator's contract is with a project promoter, then it is with this organisation that their legal responsibilities lie. They must, of course, recognise the context in which the project has been funded, but they owe their loyalties to the project, not to the paymasters nor directly to the beneficiaries (see also Section 11.1 ethics) although if evaluators feel that the project is not meeting its obligations to either of these groups they should make their feelings known to you as the project manager and possibly to the promoters (depending on their terms of reference and contract). One of the evaluator's role s is to identify problems and provide you with information *which will help you solve those problems*. Evaluators are not policemen! However, it goes without saying that if an evaluator believes that a criminal offence has been committed, they have the same legal responsibilities as any other citizen to report it to the appropriate authorities.

Under EU law, this includes misappropriation or misuse of EU funds.

### Projects are often managed by people who have little or nothing to do with the application (11)

Because of the long lead in time to many European programmes it quite often happens that by the time the project comes to fruition many of the people involved at the design stage will have moved on. The project manager, the

The period between funding application and the start of the project is often lengthy. An ex ante evaluation is useful to reestablish the objectives and targets of your project.

staff and the evaluator are not appointed until funding has been approved and so evaluators frequently work without a sense of the history of the project. Where has it come from? What experiences have led this organisation to seek support for such a project? Are the issues the same now as they were then? Have other ways of meeting the original needs been met?

The headache for evaluators is whether they are evaluating the project as conceived at the time of application or as the project ought to be now. In an ideal world, you, as the project manager, should notify your Technical Assistance Office if the project environment has changed and the project needs to respond differently. In practice, there is a disincentive to do this as it often causes further delays. For this reason, especially if the project is large, one useful piece of work for the evaluator is to undertake an ex-ante evaluation to establish what has changed since the initial needs assessment (or whatever drove the application) and whether the objectives and targets are still relevant and appropriate. At the very least you need a mid-term evaluation if, for example, the halfway point of a 3 year project is more than about 3 years away from its conception. This gives you chance to re-align your strategies and priorities to reflect any new or emerging issues.

# Chapter 5
# Who is involved in evaluation?

This section covers three main topics. Firstly we look at who the major stakeholders are in a project and how they may be involved or implicated in the evaluation process. Secondly we look at the roles and responsibilities of an evaluator. Thirdly we look at the different types of evaluation team that you can assemble and the differences between internal and external evaluation. Finally we look at the how you go about finding an evaluator

## 5.1 Who will be involved?

There can be many different stakeholders in a project but by and large they fall into 3 groups.

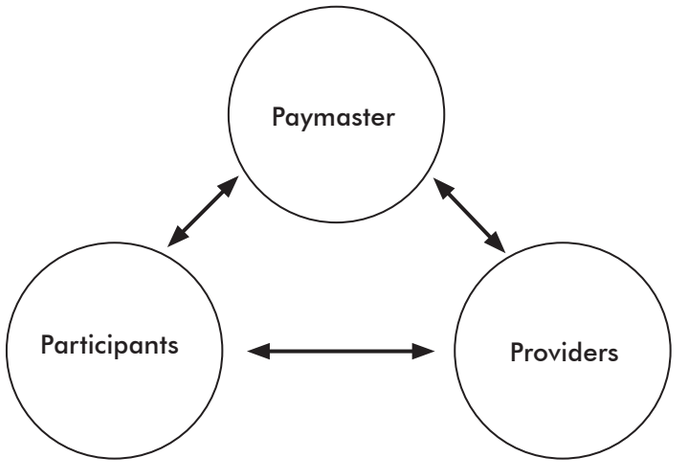


*Figure 1: Stakeholders I*

– 'Paymasters' may include the sponsoring institution, one or more external funding agencies, an executive body or an intermediary organisation.
– 'Providers' could include internal project staff, trainers, external consultants, sub contractors, transnational partners – each of whom has a partial accountability for the delivery of the project.
– 'Participants' defines all those who may be direct beneficiaries or end users of the particular project and may be individuals or groups or whole communities (both physical communities and communities of practice or research communities). The parameter is that this group is not directly responsible or managerially or financially accountable for the project – although they may have personal or collective responsibility for some aspects of it (for example, students being 'responsible' for their own learning).

Each one of those groups has something to contribute to the evaluation process (ideas, data, interpretations) and each group can use its outputs (for learning, future planning, decision making). In some ways evaluation can be thought of as a way of facilitating exchange between these groups and there is

a type of evaluation (based on social learning models), which uses that idea in its theoretical approach and in its practice (details in the 'Theory' chapter).

Remember, there may well be more than one group of people in each category and occasionally some groups fall into more than one category

It is a useful exercise for you and your project staff, together with your evaluators, to identify clearly who the specific individuals and groups are in each of those categories for your project. Ideally, all of the groups could provide data and receive feedback from each of the others. Pragmatics and resources usually make this impossible. It is, therefore, essential at an early stage that evaluators and project managers identify the most important groups in each category and clarify the key information providers and key information recipients. Then you can make a list of those that could potentially be involved in the evaluation process, what they could contribute in the way of input and what they could get out of it.

Read more on this under 'Reporting Outcomes'.

Failure to do this can result in ambiguity about both the organisational and structural 'level' of the evaluation (i.e. the level of detail or disaggregation needed) and also the form and content of the evaluation deliverables will depend on their audience. (More on this under 'Reporting Outcomes')

You may want to use a different way of sorting the groups. The diagram below may be more appropriate, for example, to research or experimental projects. Or you could combine the two. Either way, there is a checklist to help you do this in the toolkit section.

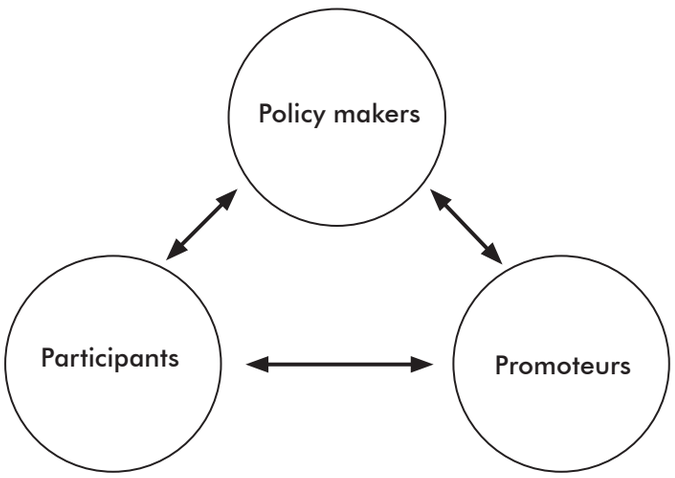


*Figure 2: Stakeholders II*

The groups in each position in the two diagrams occupy roughly analogous spaces. For example, policy makers are often paymasters and promoters will usually be the providers in a training or research project although they may be at different hierarchical levels in the organisation.

- 'Policy makers' or policy influencers may operate at a transnational, national, regional, local or institutional level.
- 'Promoters' implies sponsoring bodies or agencies, who may be involved in the future in similar activities.
- 'Practitioners' embraces the wider professional community of those directly involved with the delivery of work in the field.

## 5.2 Roles and responsibilities of the evaluator

We have already considered (in Getting Started) that an evaluator may adopt a variety of different roles – mentor, coach, consultant, provider of feedback, critical friend, facilitator and so on. All of these are legitimate roles and a good evaluator will have skills based on all of these. Also, different theoretical approaches demand different types of intervention from the evaluator and different skill sets. However, there are some core roles and responsibilities common to all project evaluators and as a project manager, this is the minimum you should expect from your evaluator.

There are core roles and responsibilities that should expected of all evaluators.

The evaluator (or evaluation team leader) should

For project manager add, in every case, 'or whoever is ultimately responsible for the evaluation'

- agree their terms of reference and 'rules of engagement'
- negotiate and agree with the project manager the questions that should be answered by the evaluation and what aspects of the project should be evaluated
- clarify the boundaries and scope of the evaluation
- be responsible for planning and organising the evaluation, sometimes in conjunction with designated others, and drawing up an evaluation plan
- be responsible for assembling and managing the evaluation team - in consultation with the project manager
- be responsible for managing the relationships involved in the evaluation process
- be responsible for designing or agreeing what indicators and criteria are to be adopted (see section on 'Indicators' and 'Criteria')
- be responsible for designing the data collection methods, although others may actually collect the data (see section on 'Data Collection')
- design or choose the data collection tools
- train others to use the tools if necessary
- manage the data processing, involving others as agreed (see section on 'Data Processing')
- suggest a range of evaluation products, agree the evaluation reporting procedures and outputs and be responsible for delivering them
- ensure that the evaluation is conducted ethically and that it is governed by a code of conduct (see section on 'Ethics')

## 5.3 What are some possible types of evaluation teams?

There are many types of evaluation teams that you could assemble. Three possible options for evaluation teams are

This section has been based on the American 'STAR' programme evaluation handbook, which is available on line at www.projectstar.org/star/index.htm

- An external evaluator (which may be an individual, research institute, or consulting firm) who serves as the team leader and is supported by in-house staff.
- An in-house evaluator who serves as the team leader and is supported by project staff and an outside consultant.
- An in-house evaluator who serves as the team leader and is supported by project staff
- An outside evaluator and an internal evaluator

Whatever team option you select, you must make sure that you, as the project manager, are part of the team. Even if your role is limited to one of overall evaluation management, you must participate in all phases of the evaluation effort.

As the project manager, it is important that you play a constant role in the evaluation.

But not all evaluations necessarily have to be objective, although this is a commonly held view and underpins the sorts of evaluation most project managers and funding agencies want. There are theoretical models (see 'Theory' section) and practices deriving from them that depend on the subjective skills of the evaluator or on the inter-subjectivity in a group.

### Option 1: An outside evaluator with support from project staff

Possible advantages:

– Because outside evaluators do not have a stake in the evaluation's findings, the results may be perceived by current or potential funders as more objective.
– Outside evaluators may have greater expertise and knowledge than in-house staff about the technical aspects involved in conducting an evaluation.
– Outside evaluators may offer a new perspective to project operations
– The evaluation may be conducted more efficiently if the evaluator is experienced.

Possible disadvantages:

– Hiring an outside evaluator can be expensive.
– Outside evaluators may not have an adequate understanding of the issues relevant to your project or target population.

Selecting this team does not mean that you or your staff need not be involved in the evaluation. You and other staff members must educate the evaluator about the project, the participants and the project environment. Other staff or advisory board members should also be involved in planning the evaluation to ensure that it addresses your project's objectives and is appropriate for the participants.

### Option 2: In-house evaluation team leader with support from project staff and an outside consultant

Possible advantages:

– An evaluation team headed by an in-house staff member may be less expensive than hiring an outside evaluator (this is not always true).
– The use of an in-house staff member as a team leader may increase the likelihood that the evaluation will be consistent with project objectives.

Possible disadvantages:

– The greater time commitment required of staff may outweigh the cost reduction of using the outside professional as a consultant instead of a team leader.
– A professional evaluator used only for consulting purposes may not give as much attention to the evaluation tasks as may be needed.
– It may be perceived as less objective than using an outside evaluator.

This second option is a good choice if you feel that you have sufficient staff resources to implement the evaluation, but need assistance with the technical aspects. An evaluation consultant, for example, may help with developing the evaluation design, conducting the data analyses, or selecting or constructing appropriate data collection tools. You will also want the consultant to help you develop the evaluation plan to ensure that it is technically correct and that what you plan to do in the evaluation will allow you to answer your evaluation questions.

### Option 3: In-house evaluation team leader with support from project and other agency staff

Possible advantages:

– An in-house evaluation team may be the least expensive option, but this is not always true.

– An in-house staff evaluation team promotes maximum involvement and participation of project staff and can be a useful form of staff development.
– It contributes to building staff expertise in evaluation, and capacity building in the organisation.

Possible disadvantages:

– An in-house team may not be sufficiently knowledgeable or experienced to design and implement the evaluation.
– Potential funders may not perceive evaluation results as objective.

If you decide on this option, keep in mind that although hiring an outside evaluator to conduct an evaluation may appear to be expensive, ultimately it may be less expensive than channeling staff resources into an evaluation that is not correctly designed or implemented.

#### Option 4: An internal and an external evaluator

Possible advantages

– This combines the advantages of all the others.
– It allows for two complementary evaluation processes, (internal and external evaluation) to be carried out at the same time.
– The evaluation is more rigorous because checks and balances are built into the structure.

Possible disadvantages

– This is the most expensive option.
– Roles and relationships can get confused.
– There may be duplication of effort if not carefully managed.

### 5.4 How can you decide what team is best for you?

Before you decide on the best team to assemble, you will need to consider two important issues.

– *The requirements of your funding agency*, which often insists that you hire an outside evaluator to conduct your evaluation. This type of evaluator is often referred to as a third-party evaluator and is someone who is not affiliated with your organisation in any way – someone with evaluation experience who will be objective when evaluating your project.
– *Project's resources and capabilities*. You can assemble different types of teams depending on your agency's resources and how you will use the findings. To determine what internal resources are available, examine, in particular, the skills and experience of your staff in planning an evaluation, in designing data collection procedures and instruments and collecting and analyzing data.
– *The information you already have available through project activities*. If, for example, you collect and regularly review information from a Management Information System (or any other organized participant database or information system) or conduct regular review sessions, entrance and exit interviews with staff and other participants or complete paperwork or logs on project progress, you may be able to use this information as evaluation data.

The checklist in the toolbox can help you decide what type of team you may need.

Whatever team you select, remember that you and your staff need to work with the evaluation team and be involved in all evaluation planning and activities. Your knowledge and experience working with project participants and the wider community are essential.

# Section 2

# Chapter 6
# Putting evaluation into practice

## 6.1 Getting the sequence right

We have identified 5 main stages in conducting an evaluation.

The sequence for your project evaluation will follow, more or less, these 5 steps.

– Framing the evaluation
– Deciding what to evaluate
– Data collection
– Making sense of the data
– Reporting

Other writers may have more or less 'stages' because they draw different cut off points - it doesn't really matter, the important thing is that the sequence is always more or less the same. The first stage is about framing the evaluation – most of which has been covered in the previous sections. By this we mean making decisions about why you want to do it (see Section 1.4), identifying the sort of project (is it about development or implementation?), deciding on the purpose of the evaluation, deciding who is going to undertake the evaluation and who it's for (see Section 3.10), gathering the necessary resources and choosing a theoretical basis or overall approach or model (see Chapter 13). The second stage is about deciding what aspects of your project you want to evaluate and what the key evaluation questions are going to be. The third stage is the process of deciding on the sort of evidence you need and gathering the data, the fourth is processing that data and drawing conclusions from it. The final stage is reporting and presenting the outcomes of the evaluation process.

## 6.2 Choosing a model

There are many different theoretical perspectives on evaluation, each one generating different models of the evaluation process, each one based on a different set of assumptions. To deal with these in detail is outside the scope of this book but for those who want to know more about the theory of evaluation, there is a brief introduction in Chapter 13.

The different theoretical perspectives on evaluation are summarised in section 13.1.

The simplest model of evaluation is based on a 'systems theory' approach and it follows the stages outlined above. This assumes that evaluation can be viewed as a simple cybernetic system that has an input, a 'translation' phase, an output and an impact. It also assumes that evaluation does not consist of a set of independent structures or functions but is an integrated process.

It is essentially about the evaluator adopting the role of an unbiased 'neutral' observer, relying on the data to 'tell the story'. The evaluator does not challenge the policy goals but simply detects 'errors' and provides objective feedback on how successful the project was, based on identifying and explaining discrepancies between the actual outputs and those that were planned. In this model the evaluation outputs or products are the most important part of the process because they provide the intelligence on which future decisions can be made. Many people assume that this is the only sort of evaluation

The 'Systems theory' approach to evaluation is a common approach whereby a 'neutral' observer (the external evaluator), through a process of data collection, translation and output provides a picture of the project's success or otherwise and, therefore, is able to inform future decision making.

- and it is certainly the most common. For this reason we have used it as a template for this handbook. However, it is NOT the only sort of evaluation and you may want to explore other options when you become more confident or if you employ an evaluator who suggests a different approach (see Chapter 13).

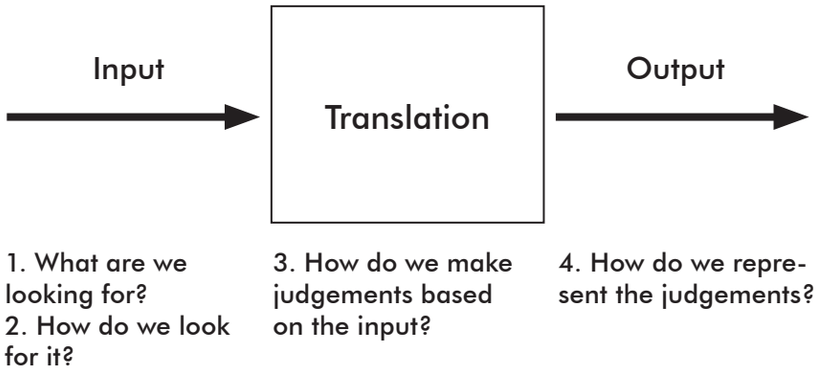


*Figure 3: A systems model of evaluation*

The Evaluation Mentor – available on line from www.evaluate-europe.net designed by the authors of this book and the CERN /EVAL team.

The model is essentially a decision making tool and can be used to design an evaluation system. The software which has been designed to support this handbook is based on this model. Each of the questions should be dealt with sequentially and each question generates a secondary list of questions or other issues around which decisions have to be made. One of the common problems in evaluation is that projects often spend too little time on the first question, rush into the second and collect vast quantities of data, miss the third (and most important stage) and then record the results descriptively without interpretation.

– *What are you looking for?* By this we mean what aspects or '**dimensions of performance**' of your project are you going to evaluate.
– *How are you going to look for it?* There are two parts to this question. Firstly, "What evidence (or **performance indicators**) are you going to use?" Secondly, "What strategies (or **data collection methods**) are you going to use to gather the evidence?"
– *How are you going to make judgements based on the input?* Again this can be broken down into a series of sub questions. How is the data to be processed? What system (or **reference base**) will be used for making comparisons? Are there quality thresholds below which performance is deemed to be unacceptable (**performance standards**)?
– *How are you going to represent the judgements?* This is simply "What will be generated as a result of the evaluation process (the **evaluation products**)?" both in the sense of a physical product or activity and also in terms of its content.

The terms highlighted in bold text are explained in more detail in the sections which follow.

### 6.3 Dimensions of performance

Often, management want to know everything about their products, services or projects. However, limited resources usually force managers to prioritise what they need to know. If the purpose of the evaluation is about development and improvement, the aspects of the project (or dimensions of perform-

ance) that you choose to evaluate should be based on what information you need to make your decisions in the future and on the resources available.

The constraints of time and money will often mean that you will not be able to evaluate every aspect of your project.

For example, you may be faced with having to make decisions because of a decrease in funding, complaints from higher up the organisation, unmet needs among staff and clients or the need to polish service delivery. Do you want to know more about what is actually going on in your project, whether your projects are meeting their goals or the impact of your projects on clients? You may want other information or a combination of these. Ultimately, it's up to you.

But the more focussed you are about what you want to examine, the more efficient you can be in your evaluation, the shorter the time it will take you and ultimately the less it will cost you (whether in your own time, the time of your employees and/or the time of a consultant).

Prioritise – be clear and focused on what you need to learn about your project.

There are trade offs, too, in the breadth and depth of information you get. The more breadth you want, usually the less depth you get (unless you have a great deal of resources to carry out the evaluation). On the other hand, if you want to examine a certain aspect of a project in great detail, you will not get as much information about other areas.

So in practical terms it is not possible to evaluate every aspect of a project, even if you had unlimited time and money. This implies a conscious selection process. It is useful to invest time in generating a list of all the possible things about your project, which it could be useful to evaluate, before making a decision. What is selected out at this stage is as important as what is selected in. We call this process identifying the 'dimensions of performance' to be evaluated. Once the 'possible' list has been produced, the following questions may help the selection process.

- Are the processes of the project to be evaluated as well as the products? In innovation projects the processes will be important, whereas in implementation projects they may not be so important.
- Will those dimensions 'selected in' provide useable and useful input in the future? Is there any potential for change? There is little point evaluating the effects of something that is unchangeable at any level.
- Who are the people at whom the evaluation products are targeted? Do these people have the ability to effect change in the areas you have selected for evaluation?
- Can the list be prioritised or weighted?
- Are the dimensions that have been 'selected in' feasible to evaluate in terms of resources, pragmatics and the timescale allowed?

As each project will be different, it is impossible to produce a useful common list of possible 'dimensions of performance' but the following ideas, whilst not in any way comprehensive, may help you start. (NB. Not in any priority order!)

| *Processes* | *Products* |
|---|---|
| internal communication | Conferences and seminars |
| external communication | training programmes |
| administration | workshops |
| finance | publications, articles, papers |
| management Information | publicity material |
| knowledge development | new services |
| industrial relations | trained people |

decision making

partnership working
health and safety
delegation
staff development
performance appraisal
recruitment and selection
reviewing
strategic planning
dissemination
public relations
legal and audit
policy making
handling meetings
recording and reporting
mentoring and coaching
line management
transnationality

websites and other computer- based items
policy recommendations
studies / audits
facilities (e.g. buildings)
strategies, action plans
networks
goods / artefacts
handbooks / manuals / guidelines
qualifications
money
case studies
structural changes
new systems / procedures
new partnerships
technology
visits
exhibitions

There are also some general questions, which are not to determine what is included in the final selection but are more about the process of selection.

– Has everyone who needs to be consulted about or involved in the selection process been included? For example, are there things the project staff want to find out more about? Do different project partners want different things?
– Do the people involved have a shared understanding and a clear definition of what is included in the dimensions of performance to be evaluated?
– Is there a mechanism for changing or modifying the list if necessary?
– Are there cut off points for these changes to take place?

### 6.4 Key evaluation questions

The questions you want your evaluation to answer will not be the same as the ones other project managers want answered. The following extract adapted from Carter McNamara provides some useful guidelines.

– If the purpose of the evaluation is developmental, what do you want to be able to decide as a result of the evaluation?
– Who are the audiences for the information generated by the evaluation (see section on stakeholders)?
– What kinds of information are needed to make the decision you need to make and/or enlighten your intended audiences, e.g. information to really understand the process of the project (its inputs, activities and outputs), the perceptions and attitudes of the beneficiaries or clients, strengths and weaknesses of the product or program, benefits to customers or clients (outcomes), how the product or program failed and why, etc.
– From what sources should the information be collected, e.g., employees, customers, clients, groups of customers or clients and employees together, project documentation, etc.
– Are there limitations on how that information be collected in a reasonable fashion? Are there some data gathering tools that would be inappropriate – if so, you need to tell your evaluation team this, e.g. questionnaires may

be of little value on a project dealing with literacy issues, observing customers or employees may be threatening to some people. This is dealt with in more detail in Section 8.5 on data collection but this question needs to be asked before the data collection stage as it may preclude certain evaluation questions being asked.

– When is the information needed (so, by when must it be collected)?
– What resources are available to collect the information?

## 6.5 A Framework

We have also found the following framework useful as a starting point, within which you can ask more focussed questions, depending on which dimensions of performance you are interested in.

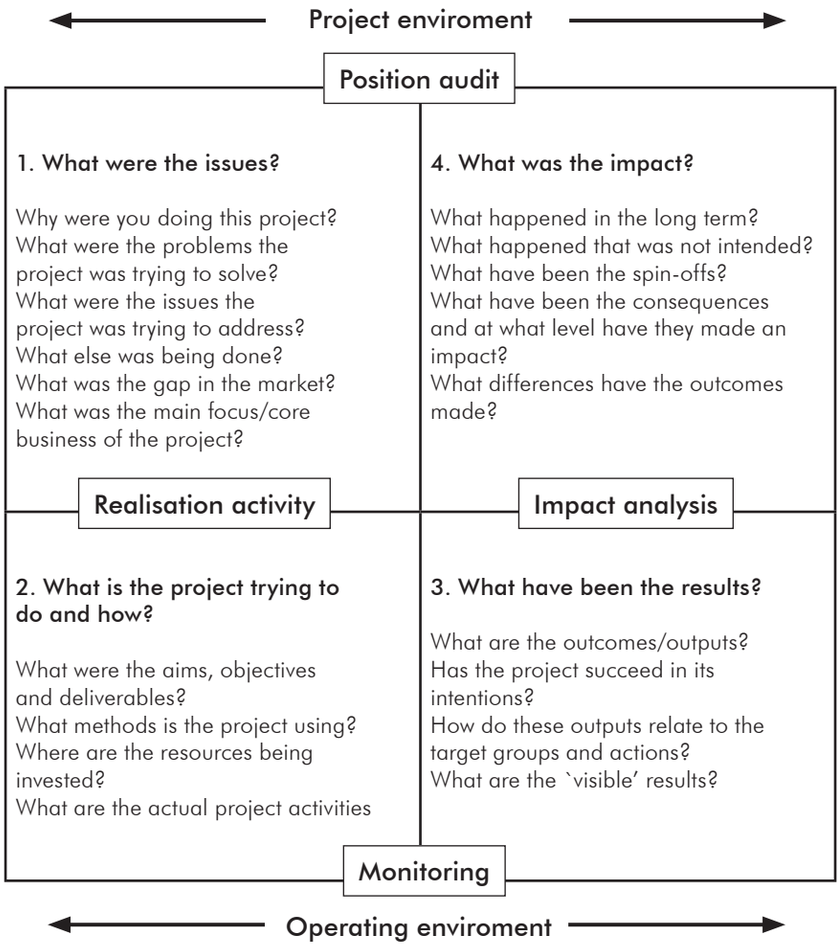


*Figure 4: The Nexus Model*

This model was developed by Nexus Research Ltd. They were one of the partners of the CERN network that produced this handbook. It is particularly helpful in providing an overview of the project lifecycle and generating the sorts of questions which the evaluation process should address at each stage.

A complete operating manual for people wanting to use this model in practice is available from the Nexus Research (www.iol.ie/nexus). There is also comprehensive computer based management and reporting software to support it.

This model assumes that there are 4 sequential stages to the evaluation process that correspond roughly to the time phases of the project. At each stage there are key areas of investigation. Examples of the sorts of questions, which the evaluator would try to answer at each stage, are shown on the

model. Crucially, the first box and the last (issues and impacts) are concerned with the project in context – where the project is located in the wider world and how it impacts on its external environment. The second and third boxes (activity and results) address the internal world of the project and how it operates.

This model can be used on its own or in conjunction with the 'systems approach' model (fig. 4), which provides a more detailed tool for constructing the actual evaluation activities within each of the 4 squares. It focuses more on the concrete operations of evaluation and is not linked to particular phases of project development.

# Chapter 7
# Performance indicators, performance criteria, performance standards

## 7.1 Do we need them?

The use of performance indicators, performance criteria and performance standards is becoming widespread and some organisations stipulate that they must be an integral part of an evaluation system. Their use is increasingly perceived as 'good practice' and may be a requirement of funding or accreditation systems. They will almost certainly be required when the purpose of the evaluation is about accountability. However, their use all too often can distort the evaluation process or at the very least predisposes to an evaluation culture that is about ticking boxes at the expense of real learning. They may also be redundant when a rigorous monitoring system is in place along side the evaluation system.

Performance indicators are often set by sponsoring organisations and/or funding agencies, but can sometimes reduce the evaluation process to a matter of ticking boxes.

We are not advocating their use nor discouraging it, only making the point that it is possible to undertake a valid and useful project evaluation without using these particular measures and some project managers will feel that they are not relevant to their particular circumstances. Either way, we feel that project managers should be familiar with the basics in order to make an informed decision!

A valid and useful project evaluation does not necessarily need to be determined by performance indicators, performance criteria and performance standards.

## 7.2 Performance indicators

In order to understand the characteristics, scope, purpose, and applicability of performance indicators (P.I.) in evaluating projects, we first need to understand what an indicator is. In general terms, an indicator is an important outcome characteristic, attribute or variable of the processes being evaluated. For example, an indicator can be as simple as an oil light on the dashboard of a car, which when lit indicates that the car needs oil, or as complex as the relationship between cost and value or the relationship between international currency exchange rates and global economic health.

Performance indicators (P.I), in terms of projects, are (usually) based on some sort of numerical 'evidence' of achievement. If any sort of mathematical analysis of data is planned, then this sort of sophistication is probably necessary and is the base line for 'hypothesis testing statistics', which is a technical area of study and which is the backbone of a lot of quantitative evaluation. However, for our purposes we are going to limit ourselves to the idea of P.I as, roughly speaking, 'counting things' to prove a point!

This is why you need to think about PI before you think about data collection methods as otherwise you run the risk of collecting a whole lot of data only to find that you have counted the wrong things.

## 7.3 Sorts of indicators

Indicators fall into three categories: process, client benefit and consequence. For a given project, a process indicator could be the number of women returners trained. A client benefit indicator could be the improvement in self confidence of the women trained or the number going into employment and

a possible consequence indicator could be the improved ratio of women to men in the local IT labour market.

Quantitative indicators refer to units of measurement.

Qualitative indicators refer to feelings and subjective judgement.

Indicators are usually quantitative but may be qualitative. Quantitative indicators refer to units of measurement. Examples include the number of learners enrolled per course, number of learners finishing a course, average pass mark of a group of students in an examination or cost per individual in terms of money expended. Qualitative indicators refer to feelings and subjective decisions and judgments. For example, if students were asked how they enjoyed the training, they may say they "loved it" or "didn't like it". If they were asked about how hard it was, they may say "too hard", "okay", or "too easy". The indicator is what you choose to gather data about. In this case it was feelings about a course and degree of difficulty.

### 7.4 Purpose and scope of indicators

Indicators can provide crucial insights for the evaluator because they can

- provide direction and focus for the evaluation. They can highlight areas of strength and weakness, which the evaluator may then want to investigate and comment.
- enable comparisons over time (backwards as well as forwards).
- enable comparisons to be made across projects (e.g. with the same objectives but different methodologies).
- be used to predict future behaviours.

### 7.5 Choosing indicators

In nearly all project evaluations, you will need more than one indicator. For example, *"the number of people accessing an advice service after a poster campaign : number accessing it beforehand"* may be a measure of the effectiveness of the campaign. 'May be' is a key phrase because it is rarely the case that one indicator can 'prove' anything. Trying to establish unequivocal causal relationships is extremely difficult in community development type projects or learning projects and you usually need several different indicators before you can draw conclusions with confidence.

There are key factors to consider when choosing your indicators, all of which may sound blindingly obvious but which are often ignored:

Read and re-read the project's funding application. Often the objectives and P.Is are summarised succinctly.

- The indicators must be based on the objectives of the project. An amazing number of evaluations use indicators that do not directly relate to the project objectives. Often this is because they have adopted indicators designed for a different context or supplied by an 'outside' body.
- The indicators used for monitoring purposes may or may not be different from the indicators used for the evaluation. For example, the evaluation brief may be to concentrate on particular dimensions of performance of the project (see Section 6.3) and ignore others. In this case not all of the overall project monitoring indicators may be relevant. Conversely, if the evaluation is looking in detail at particular aspects of the project, you may need more finally tuned indicators or may need to generate new ones. This is particularly true if the evaluation is looking at project processes as well as outcomes. We have often been asked to look at the effectiveness of a project's meetings or communication as part of an evaluation brief in which case we have set up indicators about, for example, the average attendance at meetings, the number of different communication media employed etc.

– Indicators determine what information needs to be gathered and should, therefore, be selected on the basis of what data can be realistically and efficiently collected – or even collected at all (one legendary Health Service evaluation decided to use "the incidence of undiagnosed Type 2 diabetes : mean daily carbohydrate intake"). Timescales are a frequent problem – particularly for some output indicators. For example if "number getting jobs after the project", frequently an indicator for a retraining programme for unemployed people, this information may be unobtainable until after the funding period of the project – and thus the evaluation).
– The indicators must be understood by stakeholders and be familiar to those intimately involved with the project. We recently saw an indicator for an EU funded learning project which was "number of spontaneous organic collective learning events involving ICTs".

## 7.6 How do you design indicators?

The most useful performance indicators are those expressed as ratios. For example, 'numbers of trainees enrolled: numbers completing the programme' or 'number of completers: numbers gaining full-time employment'. However, sometimes it is sufficient just to count things and what you count will be dictated by the purpose of the project. The 'performance dimensions' listed above could be used as a starting point. So for example, if the objective of a community development project was to provide an enhanced range of support services to particular groups in the community then you could count the number of services, the number of people accessing the service(s), the average number of times the service is accessed by an individual, the number of referrals, the number of 'problems' dealt with, the average duration of contact etc.

As we have already said, performance indicators will need to be designed for a specific situation. However, the following list, loosely based on one from a Web site called the Training Zone.com, may be useful – at leas it may stimulate some ideas.



*Numbers*

*....of enquiries....of enrolments....of completers....of placements....of business start ups....of trainees progressing....of jobs created....of advice and information requests....of partners/members in a network....of presentations made....of volunteers....of contracts....of publications....of placement breakdowns....of reported problems....of press coverage....of visits....of new projects....of training events and participants....of hours worked per week....of joint pieces of work with other agencies....new products....new processes....hits on a web site....*

*Ratios*

*....enquiries to enrolment....staff to trainees....completers to starters....numbers judging programmes as 'satisfactory or above' to numbers judging it 'below satisfactory'.... business plans to actual start-ups....number of contacts before project to number after project....cost per marketing outlet to number of enquiries received....*

*Properties*

*....information and advice given by nature....enquiries by type....enrolments by type....(name area....by age....gender....ethnic origin) of funding by source....of funding expenditure by category....of training by topic....by delivery method.... of staff time spent on different tasks....of promotional activities by type....of re-*

*cruitment sources....of satisfaction levels....of trainees by eventual destination....of management committee composition....of turnover....*
*Time*
*....allocation of staff time to different tasks....time allocation to particular types of trainees or clients....time given to management tasks....time given to administration....time spent on strategic thinking and planning....time spent at meetings.... time scale for completion of project stages....actual against predicted timetables.... stages at which slippage was greatest....*
*Views and opinions*
*....of enquiries....of trainees....of staff....of general public....of end users....of researchers....of politicians....of funders....of management committees....of trade unions....of other agencies....of sponsors....about the project....about the training programme....about the meetings....about the standards of service....about their future needs....about the benefits....about what should have been done....about what should not have been done....about the barriers....about the best and worse features....about the image....about motivation....*
*Compliments and complaints*
*....number and nature of positive feedback comments....number and nature of negative feedback comments....compliance with customer service standards....levels of use of complaints procedure....satisfaction levels following complaint....*
*Achievements*
*....Attainment of objectives....delivery of stated outputs....completion of innovative project....accomplishment of planned tasks.... attainment of awards / certification....existence of evidence....of case studies....of problems and solutions....of diaries and log books....of displays....of user comment....of materials produced....of needs identified....of new practices generated....of new policy recommendations.*

### 7.7 Performance criteria and performance standards

These two terms mean quite different things. One means the minimum amount of a measured indicator required for the programme or project to be judged a success. The other means the conditions or parameters that define a particular indicator. Unhelpfully, the same distinction exists but the meanings are reversed in American usage and UK usage – unless, of course, the particular evaluator has based his/her work on American literature sources! It really does not matter as long as you understand the distinction and are prepared for inconsistencies. Whilst we are more familiar with the UK convention, we are proposing to use the American terminology because it is becoming more widespread in Europe and because of the comparative dominance of US authors in terms of books on 'Program Evaluation' (which of course we would call project evaluation!).

### 7.8 Performance criteria

So, for our purposes, the term performance criteria refers to the minimum amount of a measured indicator required for the programme or project to be judged a success If an evaluator only collects the evidence and discusses it, he or she is merely describing what has been accomplished and there is little basis for rational decision-making about the success or failure of a project. However, when there are criteria, against which the evidence can be compared, then judgments can be made or conclusions can be drawn. Crudely, a performance indicator says what sort of evidence and a performance criterion says how much of it is needed!

A criterion is a rule, standard, norm, condition, or behaviour that is considered to be "good" or "ideal". Our society has thousands of criteria that govern our daily life. For example, one is to drive (in the UK) at no more than 70 miles per hour; one must get the recommended daily allowance of so many vitamins and minerals for good nutrition; one must arrive at work when the office opens at 8 a.m. Project managers operate against a background of a huge range of informal criteria encompassing widely differing descriptions or images or beliefs of what a 'valuable', 'appropriate', 'competent', 'high-quality', 'effective', 'efficient' project is like. This is a very good reason for establishing performance criteria at the outset and making these public.

Establishing and publishing performance criteria for your project at its outset is good practice.

## 7.9 How to Establish Criteria

When evaluating projects, you will soon find out there is no definitive list - not yet at least - of ready-made criteria. Therefore, project managers and evaluators must select or develop their own.

Ryans (1957) names at least three ways to develop criteria:

– the Armchair Approach, in which the evaluator selects criteria off the top of his or her head, perhaps with little forethought and no frame of reference;
– the Empirical Approach, in which the evaluator relies on research findings or experience to assist in selecting and developing criteria;
– the Rational Approach, in which the evaluator systematically analyses the situation to develop criteria based on the best information available.

To this we would add a fourth, the 'From Heaven Approach' which is when criteria are imposed from 'on high', typically by a funding agency or senior management or by some quality assurance organisation

## 7.10 Useful concepts

There are several concepts or notions you and your evaluator can draw on to generate criteria to use in measuring projects.

We have used this term to cover mean, median and modal values – all these measures may be useful in different contexts.

– Parity exists between a population and a subset of the population when their respective compositions match proportionally. For example, if Black Minority Ethnic (BME) groups make up 20 percent of the population in a community, then the potential number of black people accessing services provided by a project is also 20% and so 20 percent minority usage could be a minimum standard against which to judge whether the project meets equality guidelines. This concept of parity, of course, can be applied to other areas.
– An average is the midway between two extremes. If regional records reveal that 50 percent of the participants on return-to-learning projects went on to further training, then this average may become a standard against which to judge a new return-to-learning project.
– Research results may provide a baseline for establishing criteria. If research reveals that success at job interviews depends on an interviewee possessing a particular set of skills and behaviours, then these findings can become criteria to help judge an employability project.
– A theory or a theoretical model may generate criterion. For example, when a new method or practice is introduced, Adoption Theory (widely used in marketing) gives some insight into the percentage of people that might adopt the new method in a given period of time. Adoption Theory suggests that about 2 percent of a population are innovators and about 13

percent are early adopters. The rest are late adopters or non-adopters. A project to encourage small and medium enterprises to introduce particular ICT technology might establish a specific objective that 10% of SME will adopt the technology in the second year of the project.

– What "should be" could be a basis for establishing criteria. When evaluators are assessing your needs, they should establish "what is" the percentage condition and "what should be". The "what should be" becomes a criterion. For example, a government policy might be that all people with disabilities should have access to education or training facilities. Your project may be to provide that in a given catchment area, where the current percentage of disabled people having access to education facilities falls short of this.

Read more in section 8.6 on norm referencing, page 53.

– Comparison with similar projects is another possible way of establishing criteria (see also Chapter 7). An evaluator who analysed the cost per contact hour of instruction for a particular training programme found it was 20 € and wants to know if this is high or low. Comparison with other training programmes run by the same agency or with similar programmes run by other agencies provides a basis for developing criteria for future use.

It is important that the project manager, the evaluator and major stakeholders agree (or are aware of) the performance criteria.

It is important that the rationale for criteria is clearly established in your mind and the evaluator's mind and that you both agree. Evaluation can be disastrous if you are both using different criteria. Moreover, if the evaluator has used a logical process for selecting criteria, you can then explain or justify the criteria to other stakeholders. If the rationale is complex, subjective or theoretical, get the evaluator to put this in writing.

### Evidence

Without evidence an evaluator is at a loss in making a decision about the impact of a project. Obtaining evidence is probably the most resource-consuming aspect of conducting an evaluation. Evidence comes in a lot of shapes and sizes and it may be difficult at times to sort out the important from the not-so-important.

### Definition of Evidence

"Evidence may be acts, words, or things that provide a sign or indication". Evidence is that which provides proof of the quality of the project. Evidence, when accumulated into a pattern, provides a picture adequate for judging the extent to which criteria have been met. Rossi (1989) refers to three kinds of data or evidence.

– Data that represent the effects of the project itself
– Data that represent the extraneous confounding factors, which are events not under control of the project
– Data that represent the design effects of the evaluation, such as errors of measurement and sampling problems.

The evaluator is mainly interested in the data that represents the effects of the project but must take the other two into account and so will have to spend some time collecting data about design effects and external factors if his/her judgments are to be reliable. The following chapter looks at ways of gathering evidence.

# Chapter 8
# Collecting your evaluation data

By now you should have set out the ground rules for your evaluation, decided who will be involved, identified what aspects of your project (dimensions of performance) you want to look at and what questions you want answered and sorted out what evidence would satisfy you (performance indicators and performance criteria).

Having got this far, you now need to plan how you are going to collect the data you need. There are three main questions to be addressed:

– What kinds of information do you need?
– Where can that information be found – i.e. what are the sources?
– How can that information best be collected within your human resources, budget and schedule constraints?

In terms of time and effort - and therefore resources - data collection is probably the part of the evaluation process where the biggest investment is made. Proper planning will ensure that this effort is not wasted and will maximise its contribution to the overall project. One of the most common problems in evaluation is actually not lack of data but too much data or the wrong sort of data or data collected under conditions which make it unreliable. These issues are critical to the success of your evaluation. The higher the quality of the information collected, the better the evaluation.

Your project evaluator will spend most of their allocated time collecting data.

This chapter will give you an overview of the main things to consider when planning and reviewing your data collection activities. It also outlines the most common methods (surveys, focus groups, interviews etc.) and looks at their relative advantages and disadvantages. We have also included 'how-to-do-it' sheets for each method in the toolbox at the end.

However, data collection is a very broad and technical subject area and we cannot hope to provide a comprehensive guide. There are many excellent books on data collection, research methods, survey methodology and so on. If you are thinking about large-scale surveys or designing questionnaires that will be used as a basis for statistical analysis, we suggest that you read some of the available literature and /or buy in specialist help. Your evaluator may or may not have the knowledge and experience to use quantitative methods. This is in no way a criticism of the evaluator but is something you should discuss with them before deciding on how the data collection is going to work.

If complex or large-scale statistical analysis is central to your evaluation, make sure that your evaluator has the knowledge and experience to design and/or use the appropriate data collection tools.

## 8.1 Identifying the data you will need to collect.

The first thing that you need to consider is what information you will actually need as part of your evaluation activity. It is obviously important that the data collected is sufficient to:

– Meet the overall objectives of the evaluation.
– Provide information of a quality and a quantity that will allow you to determine whether identified performance indicators have been met or not.

The worksheet provided overleaf gives a framework to use when planning the collection of your data.

| Objective or PI | Data Elements. | Data Source(s) | Collection Method | Who | Start | End |
|---|---|---|---|---|---|---|
| Describe your first Dimension of Performance here together with any performance indicators | List the Associated Data Elements here. | List the possible sources for that data here. | What method or methods might be used to collect this data | Person | Date | Date |
| Describe your next Dimension of Performance here together with any performance indicators | List the Associated Data Elements here. | List the possible sources for that data here. | What method or methods might be used to collect this data | Person | Date | Date |

Using the worksheet, fill in your first dimension of performance and performance indicators if you are using, them in column 1. These may be expressed as broad dimensions of performance, project implementation or participant outcome objectives or they may be more general statements to do with project processes. Many authors are quite prescriptive about this and insist that the starting point for should objectives stated in measurable terms. This is O.K. if you are undertaking a traditional, straight forward 'outcome evaluation' but a bit restrictive if you are experimenting with other sorts of evaluation. The whole point about doing this exercise is simply to determine the kinds of information you need, to make sure you don't forget anything and, most importantly, avoid the problem of collecting more information than is actually required.

The above exer-cise will help you to determine the sort and the amount of infor-mation that you'll need to collect.

It is only necessary to list the objectives and performance indicators that are the focus of the evaluation you are conducting. This sounds rather obvious, but too much data of the wrong sort is one of the most common evaluation problems.

For example, you may be running a community development project and want to find out whether your marketing activities are effective or not. You may want to set a performance indicator that would count the number of new clients accessing your development project as a percentage of the leaflets distributed. You may want to know this in order to decide whether to discontinue leafleting and invest your resources in other sorts of publicity. In this case, it is important only to collect data relevant to this part of your work.

Next, complete Column 2 by specifying the information that addresses each objective or performance indicator. This information is sometimes referred to as the 'data elements'.

The data elements in the above case of the community development project are:

– The numbers of leaflets distributed.
– The number of new clients accessing your project since leaflet distribution began.

As a control, you might also want to count the numbers coming *before* the leafleting, if this data is available. This is why the earlier the evaluation schedule is drawn up the better.

What you don't actually need is the number of those new clients who heard about the project via the leaflets, or where the leaflet was a key influencer in their decision to participate. However, a lot of project managers or evaluators will automatically assume that they have to invent some sort of questionnaire or instrument that will generate this information. (This assumes, of course, that you are not introducing several new marketing initiatives at once!)

### 8.2 Identifying Potential Data Sources.

Column 3 can be used to identify appropriate sources for specific evaluation data. For every data element, there may be a range of potential sources. Some of these will be *secondary data sources*, some will be *primary data sources*.

### 8.3 Secondary data

Secondary data is basically existing data that has already been collected by someone else. For example:

– Local, regional or national databases.

– Previous surveys and studies
– Journal articles
– Published reports
– Records from other agencies

Secondary data is generally quicker and cheaper for your evaluator to collect. However, it is unlikely to provide the exact data that you need to evaluate the performance of your project.

Secondary data has both advantages and disadvantages. It is much quicker and cheaper than data you have to collect yourself and may be much more comprehensive. National databases will provide you with the data to make comparisons that you could otherwise not make, existing surveys and studies may have used sources that you could not access and reports and journal articles may give you perspectives and highlight issues that you may not have considered. However, the downside is that you may not know how valid and reliable the data is or the precise data you want may not be available.

### 8.4 Primary data

The clear advantage of collect-ing primary data is that you can collect precisely the information that you need. However, the cost will be greater than collecting secondary data.

Primary data is that which you gather yourself at source. This is time consuming and expensive but the obvious advantage is that you can collect exactly what you want, in the way that you want it. You will also know the 'data history' – the conditions under which it was collected, when and by whom – and thus any problems with its reliability and validity.

Primary data sources include

– Project records (case records, registration records, academic records, and other information).
– Project management information systems.
– Project reports and documents.
– Project staff.
– Project participants
– Members of a control or comparison group.
– Staff of collaborating agencies.
– Community leaders (in the literal sense of community and also community of practice)
– Outside experts.
– The general public.

As the project manager, make sure that you are involved as closely as possible in the decision making proc-ess that dictates the best sources for primary data collection. If you are using an external evaluator, do not absolve yourself of this responsibility or agree without consid-eration their proposal for primary data collec-tion.

In deciding the best sources for information, your evaluation team will need to answer the following questions:

– What source is likely to provide the most accurate information?
– How accurate do you need the information to be?
– What source is the least costly or time consuming?
– Will collecting information from a particular source pose an excessive burden on that person?
– Do you have access to the sources? (This could be that some documents are restricted or confidential, it could be that an individual you would ideally like to interview is too busy or does not see your evaluation as important to them)
– Do you have the skills to access the sources? (For example, particular target groups may need interviewers with skills or qualities over and above straight interview skills.)

The accuracy – cost equation is the most important decision. For example, it may be less costly or time consuming to get information about services from interviews with project staff but they may not provide as accurate information about services as you could get from case records or project logs. Conversely, there may be occasions when it is cheaper to use written sources rather

than pay for interviewer time but there may be 'hidden' stories behind the data. There is also an issue around 'spurious accuracy', when either the level of detail occludes the 'real' picture or else is a nonsensical bit of information which bears little relationship to reality or where meaningless mathematical averaging has been carried out.

A common error in data collection and data processing is using different 'levels' of accuracy at different stages in the evaluation process. (Statisticians would call this different levels of significance - so be careful when saying such-and-such a finding is 'significant') Very often people assume that if one set of data is known to be less than perfectly accurate, they can somehow 'compensate' by tightening up on the accuracy of the rest of the data. This is fallacious and pointless - any chain is only as strong as its weakest link. Or to put it another way, your *least* accurate data set will determine the overall accuracy of your evaluation.

Work with your evaluator to en-sure that the process of collecting and interpreting data is accurate and consistent. Your least accurate data will determine the overall accuracy of your evaluation.

Once you have identified the most appropriate sources for the information you need to collect, you should begin to look at how that information can be extracted or collected.

### 8.5 Data Collection Methods and Instruments.

For each data element, you can fill in Column 4 of the planning datasheet to indicate the method that you will use to collect that data or information. The main types of data collection methods and instruments you will use are:

Your project evaluation may use one, a number or all of these methods of data collection.

- Document search and review.
- Written surveys or questionnaires.
- Oral interviews with individuals (either face-to-face or on the telephone).
- Focus group interviews (either structured or unstructured).
- Extraction forms to be used for written records (such as case records or existing databases).
- Observation forms or checklists to be used to assess participants' or staff members' behaviours.
- Case Studies.

There is also the obvious use of direct observation by the evaluator – if a project says it has produced particular deliverables, can you see them, hold them, use them and so on.

The table shown overleaf gives an overview of the advantages and disadvantages of each type of instrument. In many cases a mixture of methods might be used. For example, if you are looking at completion rates for a particular educational course a document review and trawl of internal records should give you figures on the completion rate and more detailed data such as completion rate by gender, age group, course subject or the students prior educational attainment. However, if you also want to discover why students do not complete their course or training, further data collection instruments will need to be used. You might like to consider holding focus groups or conducting interviews with staff and participants. Generally speaking, the types of instruments selected should be guided by your data elements

| Method | Overall Purpose | Advantages | Challenges |
|---|---|---|---|
| **Questionnaires, Surveys, Checklists.** | When you need to get information quickly and/or easily from a large sample of people in a non-threatening manner. | – can be completed anonymously.<br>– inexpensive to administer.<br>– easy to compare and analyse.<br>– can capture large sample and get lots of data.<br>– many sample questionnaires already exist. | – might not get careful feedback. difficult to design.<br>– impersonal.<br>– in surveys, may need sampling expert.<br>– doesn't get full story.<br>– inflexible. |
| **Interviews.** | When you want to understand fully someone's impressions or experiences or learn more about their answers to questionnaires. | – can cover full range and depth of information<br>– develops relationship with client<br>– can be flexible, depending on client | – time consuming.<br>– can be hard to analyse and compare.<br>– can be costly.<br>– interviewer can bias client's responses. |
| **Documentation review.** | When you want an impression of how a project operates without interrupting the project; Consists of a review of applications, finances, memos, minutes, etc. | – get comprehensive and historical information.<br>– doesn't interrupt project or client's routine in project.<br>– information already exists.<br>– few biases about information. | – often time consuming.<br>– info may be incomplete.<br>– need to be quite clear about what looking for.<br>– not a flexible means to get data; data restricted to what already exists. |
| **Observation.** | Use this to gather accurate information about how a project actually operates, particularly about the processes within it. | – view operations of a project as they are actually occurring.<br>– can adapt to events as they occur. | – can be difficult to interpret seen behaviour.<br>– can be complex to categorise observations.<br>– can influence behaviours of project participants.<br>– can be expensive. |
| **Focus groups.** | Use focus groups to explore a topic in depth through group discussion, e.g., about reactions to an experience or suggestion, understanding common complaints, etc.; useful in evaluation and marketing | – get common impressions quickly and reliably<br>– can be an efficient way to get much range and depth of information in a short time.<br>– can convey key information about projects. | – can be hard to analyse responses.<br>– need good facilitator for safety and closure.<br>– difficult to schedule 6-8 people together. |
| **Case studies.** | Use case studies to understand situations in depth or depict client's experiences in a project, and conduct comprehensive examination through cross comparison of cases. | – Depicts fully client's experience in project input, process and results.<br>– powerful means to portray project to outsiders. | – usually quite time consuming to collect, organise and describe.<br>– represents depth of information, rather than breadth. |

## 8.6 Designing data collection tools

Once you have decided on the sort of data collection method(s) you then need to design (or customize) the actual instrument to obtain that specific information from that specific source. (As we shall see in Section 8.9 'Reviewing your Data Collection Plan', the type of data instrument used and the details of its use are also constrained by a number of key considerations.)

Read more on page 55, "Reviewing your Data Collection Plan".

If your evaluation team decides to use questionnaires or assessment inventories to collect information, for example on participant outcomes or project processes, you may have the option of selecting existing instruments or developing your own. It is not possible to identify specific instruments or inventories in this handbook as particularly noteworthy or useful, because the usefulness of an instrument depends to a large extent on the nature of your project and your objectives. However, some examples of existing instruments are provided in the toolkit.

The toolkit at the end of this handbook contains some examples of data collection instruments, that is, questionnaires and assessment inventories.

There are advantages and disadvantages to using existing instruments. The primary *advantages* of using existing instruments or inventories are:

– *They are often, but not always, standardized.* This means that the instrument has been administered to a very large population and the scores have been 'normed' for that population. When an instrument has been 'normed', it means that a specified range of scores is considered 'normal', whereas scores in another range are considered 'non-normal'. Non-normal scores on instruments assessing, for example, intelligence, substance use, self esteem, employability, recidivism and the like may be indicators of potential behaviours or events.
– *They have usually, but not always, been established as valid and reliable.* An instrument is valid if it measures what it is supposed to measure. It is reliable if individuals' responses to the instrument are consistent over time or within the instrument. The issues of reliability and validity are discussed in more detail in Section 8.10 below.

The primary *disadvantages* of using existing instruments are:

– *They are not always appropriate for all cultural or ethnic populations.* Scores that are 'normed' on one cultural group may not reflect the norm of members of another cultural group. Translating the instrument into another language is not sufficient to make it culturally appropriate. The items and scoring system must reflect the norms, values, and traditions of the given cultural group.
– *They may not be appropriate to your project.* Your objectives and the interventions you developed to attain those objectives may not match what is being assessed by a standardized instrument. For example, if you want to evaluate the effects that a counselling project has on homeless young people, an instrument measuring depression may not be useful although depression may be either a cause or an effect of the homelessness.

An existing stan-dardised data collection tool may not be appropriate for your client group or for attaining data on the precise objectives of your project.

If an outside evaluator selects an instrument for your project evaluation, make sure that you and other members of the evaluation team review each item on the instrument to ensure that the information it asks for is consistent with your knowledge about your project, the context in which it operates and your expectations about how project participants will change or behave.

Make sure that the data collection tools that your evaluator employs are appropriate for your project. As the person who knows the project best, appropriate tools can only be developed with your assistance.

If your evaluation team is unable to find an appropriate existing instrument they will need to develop one. Again, if there is no one on your team who has expertise in developing assessment instruments, you will need the assistance of an outside consultant for this task.

### 8.7 Data collection procedures.

The evaluation team will need to establish a set of procedures to ensure that the information will be collected in a consistent and systematic manner. Information collection procedures should include:

- *When the information will be collected.* This means at what stage(s) in the project will data be collected but could also specify at what point in a day or event or training programme. For example, if you are issuing an evaluation questionnaire to participants on a training course, the responses may be very different if people are asked to fill them in before they go home or complete and return them within say two weeks.
- *Where the information will be collected.* This is particularly relevant when information is to be collected from project participants and it is a good idea to be consistent. For example, participants may provide different responses in their own work environments than they would in an agency office setting.
- *Who will collect the information.* In some situations, you will need to be sure that information collectors meet certain criteria. For example, they may need to be familiar with the culture or the language of the individuals they are interviewing or observing. Administering some instruments also may require that the collector has experience or training to use a particular tool.
- *How the information will be collected.* This refers to procedures for administering the instruments. Will they be administered as a group or individually? Will people have chance to 'compare notes'? If you are collecting information from children, will other family members or a teacher be present? If you are collecting information from individuals with a low level of literacy, will the data collectors read the items to them? What level of confidentiality /anonymity will people be promised?

How you collect the data is as important as the sort of data you are collecting. You should consider the issues of when, where, who and how detailed in this section (Data Collection Procedures).

### 8.8 Training the data collectors.

If, in addition to the evaluator, you are going to be using a team of people to people to collect data – for example for interviewing or carrying out surveys – everyone involved in collecting evaluation information must be trained in the specific data collection procedures you plan to use. This could be very brief but should include:

- An item-by-item review of each of the instruments to be used in data collection, including a discussion of the meaning of each item, why it was included in the instrument, and how it is to be completed
- A review of all instructions on administering or using the instruments, including instructions to the respondents
- A discussion of potential problems that may arise in administering the instrument, including procedures for resolving the problems
- A practice session during which data collection staff administer the instrument to one another, use it to extract information from existing project records or complete it themselves, if it is a written questionnaire
- A discussion of respondent confidentiality, including administering an informed consent form, answering respondents' questions about confidentiality, keeping completed instruments in a safe place and procedures for submitting instruments to the appropriate person

It is important that the people carrying out the data collection are familiar with the data collection tools and confident in their use.

– A discussion of the need for frequent reviews and checks of the data and for meetings of data collectors to ensure data collection continues to be consistent.

It is useful to develop a manual that describes precisely what is expected in the information collection process. This will be a handy reference for data collection staff and will be useful for new staff who were hired after the initial evaluation training occurred. The development of such a manual is of particular importance where there is a large number of data collectors or where data is being collected over a long time period.

The development of a manual that describes how to conduct the data collection would be useful as a guide and reminder for existing project staff, as a training resource for future staff and as an adapt-able resource for future projects.

## 8.9 Reviewing your data collection plan.

Once you have identified your data collection needs, the instruments you will use and the procedures governing their implementation it may be useful to review progress to date before testing your instruments and beginning the collection of real data for your evaluation.

In reviewing your data collection plan and instruments there are three key considerations which should be kept in mind. These are:

– *Technical adequacy:* reliability, validity, freedom from bias, etc.
– *Practicality:* cost, political consequences, duration, personnel needs, etc.
– *Ethics:* protection of human rights, privacy, legality, confidentiality etc.

## 8.10 Reliability and validity.

It is generally agreed that 'good' measures must be reliable and valid. *Reliability* is usually concerned with stability over time and across different persons collecting the data. *Validity* is concerned with whether or not the item actually gathers the information you intended. Understanding the difference between these two terms is important.

There is a distinction between reliable data collection and valid data collection.

A reliable data collection instrument consistently conveys the same meaning. Will a person reading a question interpret it the same way each time he or she reads it? Will different people interpret it differently? If the question does not convey a single meaning, we cannot be sure which meaning the respondent had in mind when answering the question. Similarly, will two different interviewers elicit the same information? Or will two people end-testing trainees after a training course assess their competence at the same level.

A reliable data collection tool should convey consistently the same meaning and should be stable over time and across different persons collect-ing the data.

If we are talking about consistency between evaluators, we call this inter-tester reliability. If we mean does the same evaluator reach the same conclusions on each occasion (for example, are they harsher in their judgements or perceptions if they are in a bad mood!), then we call this intra-tester reliability. If the issue is consistency despite any environmental or operating conditions we call this extra-tester reliability.

Data collection instruments are said to be valid if they really 'test' what they set out to test! There are three types of validity.

A valid data collection tool is simply one that collects effectively the information you set out to collect.

*Face validity*. Is the method of data collection congruent with what is being assessed or evaluated? For example, if the purpose of an employment training project was to teach trainees to develop job interview skills, asking them to write an essay on this would not be as valid as watching them doing this in a real or simulated environment. Similarly, if you are evaluating the effectiveness of transnational meetings, the promoters minutes are a less valid measure than actually attending the meeting and observing or talking to the participants.

*Content validity.* This simply means, does the tool or method you are using actually measure or test what you set out to test. So, if you are evaluating the effectiveness of a project's dissemination strategy and simply looked at the publicity materials produced, you may actually be measuring marketing or advertising rather than real dissemination. Or if you were evaluating a project which was designed to give unemployed people skills to improve their employability, your methods would have to include ways of checking up how many actually got jobs after the project if the evaluation was to have content validity.

*Predictive validity.* Sometimes it is not actually possible to test or evaluate in a live situation what the project objectives are designed to achieve. A classic example would be training pilots to cope with crash landing situations or training people on a first aid course to cope with emergencies. It is unrealistic to test the effectiveness of the training in a real situation so we use simulations to predict how people might react. Similarly you might want to evaluate a project's sustainability but do not have the funding to undertake a long term impact analysis. This means you have to find tools or methods which are predictors of what is likely to happen.

Ideally, the data that you collect should be both reliable and valid. However, this is often an difficult balancing act.

In an ideal world, we would like all the data we collect to be both valid and reliable. However, in practice it is a bit like a see-saw – as one goes up, the other comes down. There is very often a trade-off between validity and reliability and a best balance must be found.

### 8.11 Bias and forcing.

When a response is affected by factors other than that which a question is designed to measure, then the response is biased. A biased response provides inaccurate information. For example, the problem of social desirability bias is a significant one. If a survey or interview question is about socially desirable or undesirable behavior or attitudes, there is a tendency for respondents to appear or act in a more socially desirable way, especially in a face to face situation, which may not reflect the true state of affairs. Other common biases are the people being interviewed wanting to 'please' the evaluator, wanting to withhold information because they are worried about the repercussions or using the opportunity to pursue their own agenda.

Also, you need to make sure that the wording of questions does not encourage or prompt the respondent toward a particular answer.

### 8.12 Practicality.

In terms of practicality, you need to consider whether:

Consider all of these practical points when designing your data collection plan.

- The planned timescales are realistic.
- You have the human resources, in terms of staff numbers and skills, to conduct the planned work.
- Data is not being collected which, while relevant, will be almost impossible to process or synthesise .
- The estimated cost of the work falls within the overall project budget.
- There are unreasonable assumptions built in to your data collection plan. For example, you may presume a high response rate to a questionnaire, (5-10% response to a postal questionnaire is good!) or your plan may be based on the assumption that a number of key policy makers will be willing to conduct in-depth face-to-face interviews.

### 8.13 Ethical considerations.

This is dealt with more fully in Chapter 11 but part of your data collection plan review will need to address the following questions:

– Confidentiality. Are confidentiality and data protection protocols being observed. Where, for example, respondants have been assured that their identity is protected what steps have you taken to ensure that this is the case?
– Legality. Have all legal requirements with regard to data protection been met.
– Have equal opportunities issues been addressed, is there bias in the sampling or data collection tools in terms of race, ethnicity, gender, sexuality, parental status, age, religion, disability etc.? Have cultural norms have been observed?

Having conducted your review, and altered your data collection instruments or plan accordingly, you can now move on to the task of actually collecting the data.

### 8.14 Monitoring data collection activities

Monitor your project's data collection activity in order to ensure consistency and completeness.

If your evaluation involves a large team or collection of primary data from a large number of sources, for example interviewing or surveying all project participants or beneficiaries then you may want to monitor the data collection process to ensure consistency. Nothing is more damaging to an evaluation effort than information collection instruments that have been incorrectly or inconsistently administered, or that are incomplete.

There are various activities that can be undertaken as part of the monitoring process.

You could think about implement-ing a routine and timeframe for the submission of completed instruments, conducting random observations of the data collection process, random checks of respondents and encouraging and supporting the staff responsible for data collection.

– *Establish a routine and timeframe for submitting completed instruments.* This may be included in your data collection manual. It is a good idea to have instruments submitted to the appropriate member of the evaluation team immediately after completion. That person can then review the instruments and make sure that they are being completed correctly. This will allow problems to be identified and resolved immediately. You may need to retrain some members of the staff responsible for data collection or have a group meeting to re-emphasize a particular procedure or activity.
– *Conduct random observations of the data collection process.* A member of the evaluation team may be assigned the responsibility of observing the data collection process at various times during the evaluation. This person, for example, may sit in on an interview session to make sure that all of the procedures are being correctly conducted.
– *Conduct random checks of respondents.* As an additional quality control measure, someone on the evaluation team may be assigned the responsibility of checking with a sample of respondents on a routine basis to determine whether the instruments were administered in the expected manner. This individual may ask respondents if they were given the informed consent form to sign (if appropriate) and if it was explained to them, where they were interviewed, whether their questions about the interview were answered, and whether they felt the attitude or demeanour of the interviewer was appropriate.
– *Keep completed interview forms in a secure place.* This will ensure that instruments are not lost and that confidentiality is maintained. Completed

data collection instruments should not be left lying around, and access to this information should be limited. You may want to consider number-coding the forms rather than using names, though keeping a secured data base that connects the names to numbers.

– *Encourage staff to view the evaluation as an important part of the project.* If project staff are given the responsibility for data collection, they will need support from you for this activity. Their first priority usually is providing services or training to participants and collecting evaluation information may not be valued. You will need to emphasize to your staff that the evaluation is part of the project and that evaluation information can help them improve their services or training to participants.

Once evaluation information is collected, you can begin to analyse it. To maximize the benefits of the evaluation to you, project staff, and project participants, this process should take place on an ongoing basis or at specified intervals during the evaluation. Information on the procedures for analysing and interpreting evaluation information are discussed in the following chapter.

# Chapter 9
# Data processing - how do you make sense of evaluation data?

For evaluation information to be useful it must be interpreted and judgements have to be made as to what it all means. In our experience, this is the evaluation stage that is most likely to be short-circuited or undervalued. So often evaluation ouputs simply record evaluation results rather than translating these results into useful knowledge. Many project managers and staff are intimidated by this activity, believing that it is best left to an expert. This is only partially true. If your evaluation team does not include someone who is experienced in analysing qualitative and quantitative evaluation data, you will need to seek the assistance of an outside evaluator for this task. However, it is important for you and all other members of the evaluation team to participate in these activities. This is the only way to ensure that the analyses will answer your evaluation questions, not the ones that an external evaluator may want to answer.

Definition of Judgment
Judgment is a natural operation involving comparison and discrimination. There are two important ingredients (1) evidence, which is the basis for knowledge, and (2) insight. Judgments have most value when knowledge and insight are high. Average quality judgment is achieved when either knowledge or insight is absent. The poorest judgment occurs when both are absent.

A dictionary de-fines 'insight' as the power or act of seeing into a situation, the act or result of apprehending the inner nature of things or of seeing intuitively.

## 9.1 Turning data into knowledge

There are basically four steps in this 'translation' process:

### Converting the raw data into collected data.

This could include statistically 'summarising' the numerical data, grouping or clustering observations and comments from interviews etc. by similarity of content, source, positive or negative responses and other factors. As far as possible it should also include a summary of 'outputs', deliverables, performance against targets and other monitoring data. The toolbox has some notes on how to do this.

Read more in Chapter 15 Toolbox.

### Checking and analysing the data

This involves checking the data in terms of its reliability, validity, conditions under which it was collected, context and so on. There may well be factors which need to be taken into account at this stage, for example, interviews undertaken at a particularly "low" or "high" point, background circumstances which could have affected the data (e.g. changes in staff). Although big problems should have been ironed out at the data collection stage, there will inevitably be some issues that will have 'slipped through the net'. You may find, for instance, that some of your data will have to be scrapped.

Ideally, the data collection process should be continually monitored for reliability and validity. However, at a minimum, the reliability and validity of your data collection process should be checked before analysis of the collected data.

### Interpreting the data

By this we mean looking for patterns, themes and explanations. It is important at this stage that precipitate conclusions are not drawn and the whole range of possible interpretations and explanations are considered. It also involves identifying probable cause and effect relationships, general trends and directions. It may also include setting up some hypotheses that can be tested and it is at this point that gaps in the data are most often identified - that is, more evidence is needed in certain areas.

### Drawing conclusions

Finally we need to draw some conclusions from the range of possible interpretations. How this is done will depend entirely on whether the purpose of the evaluation is developmental or to do with project justifications. What follows is largely concerned with evaluation for improvement although the section on Performance Standards and benchmarking has been included which will be of more use to project accountability-type evaluation.

## 9.2 What does the data mean? Outcomes and results

It is almost impossible to provide comprehensive guidance on how you can make decisions about what your data means as obviously each project will be very different - different objectives, different data, different results. However, the following ideas may get you started.

You can:

- *Describe* what you did (or are doing), who did it, and the characteristics and number of participants. You can then *compare* this information with your initial plan and determine whether there is a difference between objectives and actual implementation.
- Identify the *reasons for differences* between your planned objectives, planned activities and expected results or targets and what you actually achieved.
- Identify *barriers encountered* and *factors that facilitated or helped*.
- Identify things the project *should have done more of*.
- Identify the things that the project *should have done less of or stopped doing at all*.

A final evaluation report will generally cover all of these points.

For example:

| Planned | Actual | Differences | Reasons | Barriers | Facilitating factors |
|---|---|---|---|---|---|
| Produce 6 training packs for community development workers | 5 produced | Shortfall of one pack but the same output in terms of content | 2 packs were combined because the material was around similar themes | Should have identified this at the planning stage – did not include the writing team early enough | The writing team were flexible in their approach and alerted the planning team early on |
| As part of an Active Citizenship programme, run 6 x 2 hour training sessions for 72 young people to encourage them to participate in local democracy | 4 x 2 hour sessions run with average of 9 young people plus 1x 1 hour session | Attendance dwindled after the third week<br><br>The last session was cut short.<br><br>The actual number of hours delivered was 9 not 12<br><br>numbers attending were 45 not 72 | The sessions were too long to hold their attention | The evening clashed with football training.<br><br>The material was not stimulating enough | The local youth club were helpful in recruiting.<br><br>The young people thought the venue was good<br><br>The young people liked the facilitators |

From these questions, you have the basis of a report. For example, this may be an extract:

"The six two-hour sessions planned were cut to five because attendance dwindled after the third week and there did not seem to be sufficient numbers to continue the programme. It is possible that we had 'saturated' the potential market or else the previous sessions had not been interesting enough to stimulate demand by word of mouth. However, the sessions were timetabled for the same night as football training and this meant the number of young people available was limited. We will make sure that if the programme runs again it does not clash with other activities.

We should have done more marketing and relied less on the youth club for recruitment and should have spent more time on preparing the materials for the sessions. However, the young people that did attend felt the facilitators were good and so was the venue.

## 9.3 Consequences and impact

The earlier questions were judging the immediate results or outcomes of the project against those intended. However, you should also include the unintentional or 'knock-on' impact on policy or practice or on individuals whether good or bad.

- The impact on individuals is called the *participant outcomes*. (See also Section 2.2) In the example above, you may be able to say that of the 45 young people who attended about half expressed an interest in participating in community activities or that 5 have joined the Community Youth Forum or whatever. A critical question is: *Did some participants change more than others and, if so, what explains this difference?*
- Again, using the same example, you might be able to say that the Community Youth Forum has now got an increased membership or a more representative membership or that the local council are considering extending the project to other areas. These are evidence of impact on practice and policy respectively.

The crucial point is that *if changes occurred, were they the result of your project's interventions?* Can you provide the evidence to back up this statement? (Using the same example, could you go and ask the new members of the Youth Forum why they had joined after attending the project) Quite often it is impossible to provide hard data but it is perfectly acceptable to use subjective judgement as long as you explain this and are cautious in the statements you make.

## 9.4 Statistical procedures

If you want an evaluation to include quantitative techniques and if you or your evaluator are experienced in handling statistical procedures, you may want to include a statistical analysis as part of your data processing. However, this is another specialist field and outside the scope of this handbook. The most important thing is to decide this in advance and to ensure that the statistical procedures are designed *before* you collect any data. We have already laboured the point that the sort of data you want to collect should be planned in advance of the actual data gathering process. If this data is to be subjected to statistical tests, you will also need to gather the data in a particular *format*, which will depend on the tests you are going to use. This should be included in the evaluation plan.

If you intend for your evaluation to contain statistical analysis, make sure that your evaluator has experience in handling relevant statistical procedures.

The following section will not turn you into an expert but may help you understand the terms your evaluator will use

## 9.5 Understanding statistical procedures.

Statistical procedures are typically used in evaluation to understand changes occurring among participants as a group and to make predictions about how they may behave or whether differences between two groups or data sets are significant or to check out whether ideas about why certain things that have happened are actually true. These all rely on very different sorts of statistics.

*Descriptive statistics* use a range of established procedures and tests to describe two or more groups (or populations) to see whether they are mathematically the same. If they are not the same, then certain conclusions can be drawn.

*Inferential statistics* make predictions or deductions - for example, about what may have been the case in the past or what may be the case in the future or whether a particular factor could have been responsible for something happening.

*Hypothesis-testing statistics* are a special category of inferential statistics that helps you to find out whether some 'hunches' you may have are likely to be true - at least, that they are not due to chance.

For instance, if we administered the same questionnaire to groups of young people drawn from different countries and different backgrounds and the results showed no statistically significant differences, then we could assume that country and background made no difference to their behaviour with respect to the issues covered by the questionnaire. If however, the tests concluded that the groups were not mathematically similar, we may conclude that country or background affected young people's behaviour with respect to particular things.

For example, in many instances, some of your project participants may change their behaviour by the end of your project. How much they change may vary; some participants may change a great deal, others may change only slightly and still others may not change or may change in an unexpected direction. Or you may want to find out whether the changes depend on how often they attend or on what day of the week or how they were recruited. All these different factors are properly called 'variables'.

### Variables

Variables can be put into different groups. Typically, an evaluator may be trying to find out whether something (e.g. trainee success) depends on something else (e.g. the quality of the teaching, the length of the programme, the recruitment policy or whatever). The issue being investigated – in the example above, this was 'student success' – is called the Dependent Variable (or DV). The factors which affect it are called the Independent Variables (IV). Most of the time evaluators also find that there are lots of other variables that they have to take into account even if they are not part of a study because they influence the results in some way. For example, the evaluator might be interested in finding out whether the effectiveness of a particular e-learning programme depends on how much teacher support is provided. S/he may want to look at a group of trainees where there was a lot of support and compare it with a group where there was very little. This is fine – until you find out that the people in one of the groups happened to be a lot older than in the other group, that some of the people were experienced computer users and others were not and that some learners were attending voluntarily and others had been sent by their companies. All these other factors may well make a difference to the effectiveness of the programme but are not actually to do with the evaluation questions you want to answer. We call these Irrelevant Variables (IR) – even if they seem very relevant! One of the skills of the evaluator is deciding which of these IR's to ignore and which to take into account. Once that decision has been made, the evaluator also has to design ways to

In American books, in particular, you may find that these IVs are called 'mediating variables'

*control* for these variables to minimise their impact on the particular relationships being investigated. There are all sorts of ways of doing this, for example by using different *sampling techniques*, using *'matched pairs'* of learners in each group, using each learner as their *'own control'* and so on.

Matched pairs means making sure that each learner (or 'subject') in the test group is matched to a corresponding one in the control group in terms of all the IVs which you think are important. That is, in the example, you would need to make sure that older learners in one group were paired with older learners in the other. Or those with the same level of computer literacy or those who volunteered as opposed to those who were conscripted were matched.

### Statistical significance

Statistical procedures are important tools for an evaluator because they can determine whether the changes demonstrated are the result of a chance occurrence or are caused by the variables being tested. This is called statistical significance. Usually, a change may be considered statistically significant (not just a chance occurrence) if the probability of its happening by chance is less than 5 in 100 cases. This would be called a 95% significance level – or, sometimes, 5 degrees of freedom. However, in some situations, evaluators may set other standards for establishing significance, depending on the nature of the project, what is being measure and the number of participants. The important thing is that they state what the significance levels are!

### Errors

Sometimes, even with valid and reliable data, the data processing will have errors in it, which will lead to the wrong conclusions being drawn. There are several types of errors. Some you are likely to make, some are uncommon.

- *Type 0 errors:* Type 0 errors are common - they occur when you fail to count or classify what should have been classified. Often called errors of omission.
- *Type I Errors:* Also quite common – type 1 errors are 'false positive errors'. That is, when a conclusion is drawn that something has happened when it has not or with insufficient evidence to substantiate that it has.
- *Type II Errors:* These are less common and simply mean that the conclusions drawn say that the results are not significant or are negative, when in fact they are positive. They are also called 'false negative' errors and occur infrequently, possibly because project managers and evaluators tend to be biased in favour of positive results!
- *Type III Errors:* Type III errors are statistically very rare and complicated to understand. Very crudely, type III errors occur when you reach conclusions that are diametrically opposite to those you should have drawn because of sloppy statistics!

Using the subject as their own control means testing the same learner under different conditions, for example, the same learner undertakes a programme with tutor support and then without.

Most of the time you will only encounter these terms in the context of statistical analyses of quantitative data. However, a rough understanding of them is helpful, particularly as some of them have spilled over into every day evaluation jargon even when talking about qualitative data.

# Chapter 10
# Evaluation products

## 10.1 Reporting results

With your evaluator, agree a format for your evaluation product(s) that is most suitable for its intended audience. Written reports are the most common evaluation product, but not the only option.

Evaluation products are simply the outputs of the evaluation process but it is useful if these are planned as early as possible in order to allow enough time and money to produce them. It is generally taken for granted that the outcomes of an evaluation process will be one or more written reports, including a 'final' report. Usually, this will be the case. However, it is not the only type of legitimate evaluation product and there are other possibilities that may be more useful, depending on who and what the evaluation is supposed to be for. Similarly, there are many different types of written evaluation reports, depending on the purpose and the intended audience. The crucial questions, as in all communication, are: What is to be communicated, to whom, by whom and when? It is useful for both the project staff and the evaluator to know at the design stage what the answers to these questions are so that the necessary resources can be allocated.

## 10.2 Alternative evaluation products

Someone once came up with the statistic that project evaluation reports are read by an average of 4.5 people and executive summaries by about 14. Whether or not this is true - or how the researcher 'tracked' the evaluation reports or in what context we have no idea. Nevertheless, in our more cynical moments we think this is quite plausible! What is undeniable is that an awful lot of evaluation reports sit unread on countless shelves across Europe and an awful lot of evaluation findings that could usefully inform future work never see the light of day. To this end, we are actively encouraging project managers to experiment with a wider range of evaluation products targeted more carefully and with a more focussed content.

In recent years, evaluators have become far more creative and advances in technology have made possible new ways of reporting on evaluation findings. These may replace or supplement traditional evaluation reports and, for those of you who are managing European projects, it is reassuring that several of the programme Technical Assistance Units are actively encouraging this as a way of making evaluation more effective, relevant and sustainable.

A written report is not always the best method of dissemination. As a supplement or as an alternative, give some thought to the broader range of evaluation products described in this section and in the Toolbox.

We have listed a full range of possibilities in the toolkit but before moving on to talk about the more orthodox type of evaluation reports we would like to highlight some of the alternative techniques we have found effective.

## 10.3 Staff development as an evaluation product

The findings of an evaluation may be presented as a staff development event – either delivered by the evaluator, if they have training skills or by someone else in consultation with the evaluator. This could be a straight forward seminar or a more interactive workshop in which participants actively work with the evaluation material in order to generate action plans for example.

Alternatively, the evaluator could generate some staff development materials based on the evaluation process and its outcomes which could be available for people to use at a later date.

### 10.4 Visual materials

Although these may incorporate text, we are suggesting that evaluation products which depend on high visual impact may be very effective in certain circumstances – particularly if you are trying to reach a large audience in a short time. We have successfully tried evaluation exhibitions which made use of annotated photographs, graphs, diagrams in exhibition display board sizes and also incorporated 3D objects. We have also presented reports in a Powerpoint format (presented by the evaluator or the client) as videos and in a 'newspaper' or magazine format.

### 10.5 E-tools

Using web based ICT could make your evaluation product(s) more interactive, inclusive and visually appealing.

The use of information and communication technologies promises new approaches to evaluation reporting. Firstly the use of web based ICT enables the evaluator to publish findings quickly and allows everyone involved in the evaluation process to become involved through on-line discussion and commentary resulting in a more interactive reporting process which encourages 'ownership' of the evaluation findings by a wider range of stakeholders.

It also allows evaluation reports to be written in 'chunks' so that different users can combine these 'chunks' in novel ways for their own purpose. At the same time the use of ICT allows a wide range of graphic representations of the evaluation products as an alternative to the usual paper and text report.

In the future, as more e-tools are developed for data gathering, it is likely that there will be greater integration between data collection and evaluation reporting together with add-on communication tools, so that evaluation and evaluation products become part of the project development cycle and communication systems, rather than stand alone outputs.

### 10.6 Writing evaluation reports

Be very clear with your evaluator about what you want from the written report(s) and who will be the audience. Over-long and unfocussed reports will go unread.

Although you may want to experiment with some of the ideas above, most project evaluations will need a written report recording what you have learned about your project from the evaluation. However, there are many different ways of reporting evaluation information, depending on how you want to use the report and who your audience will be. You need to be very clear about this when you brief your evaluator. An awful lot of evaluation reports try to be all things to all people and end up being useful to none of them. Over-long, unfocussed, glossy publications which are little more than a narrative account of the project usually end up unread on bookshelves or in bins but represent a considerable investment of evaluation resources. In this section, we suggest ways of preparing evaluation reports that are appropriate for a range of uses.

For example, an evaluation report can

- Provide a narrative of the project, identifying what it set out to do, how it was implemented and what it achieved.
- Provide evidence for existing and potential funding agencies that the project was effective and represented value for money.
- Inform management decisions by identifying changes that may be needed in the future
- Inform future policy decisions

Who is going to read the evaluation report? Make sure that you and your evaluator know from the outset.

– Contribute to the wider community of practice by expanding the knowledge base.

A list of potential audiences is included as part of the toolkit. They include paymasters (such as funding organisations or organisational executive boards), providing institutions such as management committees, senior managers, project staff, tutors etc. and also the wider public including other practitioners and researchers. The critical thing is that evaluation reports should be interpretive rather than just descriptive and, although most project managers would prefer a 'good news' report, they should reflect the problems and unresolved issues as well as the successes.

The evaluation conclusions can be expressed in a variety of ways:

– Lessons learned about the activities (content) and also the processes (form) of the project.
– Recommendations for future action for policy makers, researchers, promoters, practitioners etc.

A useful evaluation report should always be interpretive rather than just descriptive.

– Prescriptions for future action for policy makers, researchers, promoters, practitioners etc.
– Suggestions for improvement.
– Identification of critical success factors and also problem areas
– Plans and also predictions for long term impact analysis.
– New insights and alternative meanings
– Contribution the project has made to the overall 'market' gap it set out to fill or to the original problem it tried to solve. (This could be framed in terms of achievements against the original aims, objectives, deliverables or the broader and longer-term impacts).
– Identifying areas of new work

One of the problems with externally funded projects, particularly those supported by the European Commission, is the pressure to succeed, which is invariably interpreted as achieving positive outputs. As so many of the European Commission programmes are for experimental and pilot projects it is highly unlikely that all will succeed. In our experience about 20-30% of projects will actually break new ground and produce long term benefits, about the same percentage will 'fail' and the rest will fall in between. That is, they will succeed at perhaps a local level or with a limited usefulness

Unfortunately, 'failure' often results in financial penalties and both project managers and evaluators are under pressure to make the project look good. As yet, there is are no platforms for learning lessons from ` failed' projects, which, in terms of learning, can contribute as much, if not more than the successful projects to the collective knowledge pool. There are many events that disseminate 'best practice' and model projects but there are no legitimate mechanisms for projects to say 'We didn't get this right - and this is why!' Interventions which do not work, despite genuine effort, can contribute positively to the knowledge pool and project methods which show promise need further exploration.

This is potentially another role for the evaluator. We are not suggesting that evaluators should 'name and shame', rather that they should be able to capture the experience of 'failed' projects and use it positively to improve the learning of other projects. It is also important to present results that may not be conclusive but show promise and warrant additional study. For example, if there seemed to be an improvement in a particular area or with particular clients as a result of particular project activities, this is worth presenting so

that future evaluation can explore this further. Currently, so little is known about what does and does not work in some areas of vocational education and training, that any information on these issues greatly increases knowledge in the field.

### Evaluation report for funding agencies

It is possible that the form and content of your evaluation report will be determined by the requirements of your funding agency.

Reports for funding agencies may either be commissioned by the agencies themselves or commissioned by projects but with their funding agencies as a target audience. As this handbook is for project managers we are only concerned with the second type of report which will almost certainly be used by your sponsors for 5 main purposes:

– To justify their initial investment.
– To secure follow-on funding or establish their confidence in your organisation so that they make further grants to new projects you may want to run.
– To inform their spending profile so that they roll-out successful projects elsewhere or mainstream the successful activities of the project in some way.
– To provide publicity for the funding agency and demonstrate to a wider audience how effective their investments are and how they support broader policy fields.

Funding agencies are frequently responsible for programmes, which in turn finance a portfolio of individual or themed projects which collectively support their policy objectives. Although programme evaluation is a separate subject, it is worthwhile bearing in mind that most programme evaluations will focus on:

– Achievement
– Relevance and timeliness
– Responsiveness
– Impact and effects

That is, they will want to know what actual project deliverables are and whether these matched expectations, whether they financed the right projects which delivered the right things at the right times, whether their commissioning strategy is flexible enough to reflect a changing environment and whether they are creating long term change. As a project manager it is worth bearing this in mind because anything your report can do to provide answers to these questions is likely to be favourably received!

It is likely that your funding agency will be evaluating the impact of the funding programme across their portfolio of projects.

Reports for funding agencies need to be brief (they will have lots of reports to read), cautious in what is being claimed and with hard evidence and a sound evaluation methodology to back it up (they will be open to public scrutiny). It is useful if the report contains a few 'sound bites', which can be used for publicity purposes, or a summary of the evaluation report written as a press release. Get a journalist, PR or press officer to do this, not the evaluator!

The report should also specify carefully the project context and the target groups and should make explicit any links between these and the outcomes. This helps to avoid 'successful' projects being replicated in other environment for which they may not have been designed. Funding agencies are also likely to be 'nearer' policy makers – or, in many cases, will be policy makers. This means that the evaluation report should contain a short section outlining

the 'transferable' outputs, generalisable 'lessons learned' together with some policy recommendations.

Finally, funding agencies will almost certainly want to know whether the project represented value for money, whether it was efficient as well as effective and whether the cost per output compared favourably with other initiatives.

A sample outlines for an evaluation report for project funding agencies is provided in the toolbox.

### Evaluation report for project management bodies

By project management bodies we mean management committees, boards of governors, steering groups, executive groups and so on and we might also include senior managers and others in your agency who have not been directly involved with the day to day running of the project.

The evaluation report will have several purposes:

– It should inform future management decisions, particularly if the project - or a similar one – is to run again.
– It will prove to stakeholders that the organisation is self-critical and transparent
– It should demonstrate how organisational plans are being put into practice
– It should provide feedback to management about the organisation's capacity to run projects
– It should identify any new gaps in provision thrown up by the project

Too many evalua-tion reports designed for management bodies dwell excessively on project context, objectives and methodology. This is information that should already be familiar.

Reports for management will probably be quite comprehensive as they will be interested in every aspect of the project – the processes as well as the outputs. However, they can afford to be less detailed about the context, the methodology and the needs and target groups the project is responding to as they will probably be familiar with these already and may well have been involved in the original design.

Evaluation reports for managers need to concentrate on how the work of the project can be improved, built on and taken forward. It may feed into budget, staffing and communication strategies so successes and problems in these areas need reporting on. Internal and external relationships and their impact on the project, partnership working and project management should also be included.

In some ways this is the most difficult of all evaluation reports for an external evaluator because it is often the management committee who are the contractors and evaluators often find themselves walking a kind of tightrope between the management and the project staff. Management often see 'their' evaluator as a sort of 'spy in the camp', reporting back on what the project workers are doing. Conversely, project staff frequently treat the evaluator as another arm of management. Neither of these positions is a healthy role for an evaluator and as a project manager you should try to ensure that your evaluator is not drawn into the internal politics of the organisation. All too often there is pressure from both sides to represent their views and opinions in a report rather than the evidence collected by the evaluation.

Whilst evaluators should avoid acting as 'advocates' for any particular group, an evaluation report can nevertheless provide information and interpretations from different perspectives, which may help to reconcile different

perceptions or even disputes. It can also highlight good practice and thus, indirectly, identify staff responsible.

Above all, an evaluation report for managers should focus on linking outcomes to project processes, establishing causal relationships between implementation and results and identifying the unexpected issues and problems and how they were resolved.

### Preparing an evaluation report for project staff

You should ensure that your project's staff are an audience for the evaluation product(s).

Providing feedback for project staff, particularly as part of a formative process, can be a crucial evaluation role. However, it is often overlooked as an evaluation product. That said, it is questionable whether a written report is the best medium for communicating with project staff who the evaluator may have been working with closely over along period of time. Feedback seminars or staff development events such as training days based on the evaluation findings and other face-to-face strategies may be more appropriate.

Project staff will not need to be acquainted or reminded of background. A far more specific report will be needed.

However, if a written report is required then there are some points worth considering. Reports for staff will almost certainly need to

- focus on specifics rather than on generalisations
- contribute to the professional development of staff
- provide feedback which is useable and practical
- concentrate on improvement and development

The staff will be intimately involved with the project so the evaluator will be able to skip a lot of the background that other audiences need. Specific problems that the project has encountered should have been dealt with during the lifecycle of the project so there is probably no need to go over the detail but to concentrate on how things could be done differently next time to avoid similar situations reoccurring.

A major problem with evaluation reports for project staff is avoiding 'naming and shaming' individuals – if not directly, then by implication. In reports for more remote audiences a degree of anonymity is relatively easy to maintain; this is more difficult with a small group. Whilst it is the evaluator's job to report accurately, the project processes, this should not spill over into performance appraisal of particular staff.

In order to avoid discord amongst your staff, agree with your evaluator an appropri-ate tone and use of language for the evaluation report.

Giving feedback to staff who you may know at a personal level has its own rules, which are rather different from those governing other sorts of evaluation. Firstly, 'own' the feedback. Particularly if the feedback is verbal, it is important for the evaluator to use the first and second person rather than to try and 'objectify' the issues. So, if talking to, or writing a report for, a project staff team, it is better and more personable to say *"I believe that the project would have benefited if you had thought more carefully about your target group before drawing up your marketing strategy..."* than *"The project would have benefited if the marketing had been more targeted"*.

The first invites a response and promotes interaction between the evaluator and staff. The second is categorical and cuts off continuing discussion.

"Something I might have done differently...." works better in a report to staff than "One of the things that was wrong with the project was..."

Try and 'sandwich' the feedback e.g. One thing you thought was good or worked well, one thing you think could have been done differently, one overall summary statement either positive or neutral. Try not to put too many negatives in a report to staff. Even if the project has few redeeming features, focus on the positive and make sure negative feedback is constructive. This is not to say the evaluator needs to 'whitewash' a bad project but if the purpose of the evaluation is developmental, overly critical or heavy handed reporting

is more likely to demoralise and demotivate the staff team and unlikely to improve performance.

Make sure the feedback focuses on things that the staff team can change. There is little point telling them that the project was under-funded if they have no control over the budget or criticising the venue if there is no other building available.

### Preparing an evaluation report for the wider public

Preparing an evaluation report for the general public is difficult. Think about being brief, using simple and attention grabbing statistics or provide a human story as an introduction to general findings.

This is a bit of a catchall as the 'wider public' could include a very disparate range of stakeholders all wanting different information.

National and local organisations working in areas which are similar to the project, other project managers, the research community, advocacy organisations or pressure groups and relevant government departments will not want to read lengthy narrative accounts of the project. The most useful way of structuring reports for these groups is probably one paragraph on the aims of the project, the target group, the context and the dates (very important for referencing but often forgotten) followed by a brief summary of 'What worked and why' and "What didn't work and why'. It could also include issues needing further investigation. In fact, it is very similar to the abstract that precedes academic papers. Most important is a contact point for further information and a list of other reports or articles written about the project.

One of the most difficult audiences to reach is the 'general public'. This could be the local community in which the project is taking place, local politicians, the beneficiaries of the project, other local agencies and so on. We are not convinced that a written evaluation report is necessarily the best option. We have tried photographic exhibitions, newspaper articles, leaflets in libraries and community centres, pages on community web sites, making presentations at public meetings and community forums as well as written reports. If a written evaluation report is needed, then it will need to be short, attention grabbing and have local interest. Statistics are a good opener *(..."one in every three women who left school at 16 are interested in returning to learning"..."50% of new computer users are over 50"..."43 young people in Blanktown take part in exchange programme"...)*. Alternatively the evaluation report could focus on tracking one or two trainees to give the report 'human interest' and then introduce more general findings. Admittedly this is more like journalism than writing an evaluation report but your evaluator should at least be versatile enough to write 'populist' reports as well as more conventional reports without losing accuracy or reliability.

### Consultation

During the reporting stage it is advisable to organise a consultation round with the main audience groups before and normally the evaluator should produce a draft report for consultation before the final version is agreed. A typical consultation would invite readers to comment on:

- Matters of fact; it may well be that particular details (dates, times, places etc.) are inaccurate and can be corrected.
- Matters of opinion; either the evaluator will change their opinion in the light of new information and argument or will maintain their original stance. In the latter case it is probably useful to let the comment stand and qualify it by saying *"in the opinion of the evaluators ... however, this is*

*disputed by many people / the majority of those consulted felt this was not the case / one person disagreed with this"* etc.

– Errors of omission. The most frequently occurring response to consultation is that the report does not contain reference to a particular aspect of the project. Ultimately this is for agreement between the evaluator and the contractor and is almost always as a result of word length rather than deliberate exclusion. However, omissions may also be for reasons of confidentiality (e.g. staffing issues), problems that have been dealt with satisfactorily and from which there are no further learning points or matters of a sensitive nature or which have a relevant audience more limited than the circulation of the document.

Ensure that key project staff and stakeholders are able to read and respond to the evaluation report(s) before they are finalised.

The important thing is to discuss authorship of documents and ownership of evaluation products and who has powers and rights of censorship before they become issues. In addition, it may be that some of these problems can be avoided by having multiple evaluation products. Chapter 11 deals with these issues in more detail.

## 10.7 Structuring Evaluation Reports

Most evaluation reports share common elements, which are outlined in the Toolkit.

Each evaluator will ultimately want to structure their final report differently. Nevertheless, there are 'elements' that appear regu-larly in final reports and we have included several proforma in the toolbox to help evaluators and project managers agree the scope and content of a final report by selecting key items. Which fea-tures are included will be determined, ultimately, by the purpose of the evaluation, the purpose of the report and the intended readership. An evaluation report also needs to be seen in the context of other project documentation and output. So, for example, if there is a narrative (non-evaluative) report on the project itself produced from another source, the evaluation report may be able to omit detailed descriptions of the project. Conversely, an evaluation report may be the only written output of a project and will thus need to be more comprehensive.

C.L.Taylor and Larry Arrington, whose work we have used exten-sively, offer the following useful tips for evaluators:

– Learn as much as possible about your audience.
– Have some understanding of the amount of knowledge the audience has of the project.
– Consider the education level of your audience. Keep in mind the cultural background of those interested in the report.
– The audience's profession or occupation is a key to enhancing understandability of the report.
– Administrators and policy makers don't have time to review volumes of material or listen to hours of oral reports. Therefore, report in small amounts and often.

### Sustainability of evaluation.

With a view to sustaining your project after the initial funding period, you could design your evaluation process to include elements of staff development, increasing your staff's ability to reflect on, review and analyse their own work.

Many projects, by their very nature, are meant to be short-term and experimental. They may be trying out new ideas and approaches or funding agencies may be pump priming development costs with a view to mainstreaming the outcomes. However, there is a tendency for projects that were conceived as pilot projects to become dependent on the original funding source and continue drawing down funding after the innovative phase. For this reason,

sustainability is becoming an increasingly important issue, par-ticularly for projects supported by EU programmes.

As evaluation is a critical process and represents a real cost to the project, we have an interest in working with clients to develop an 'evaluation culture' within their organisation and to encourage organisational staff to build their in-house capacity to undertake their own evaluation. Whilst some organisations in some circum-stances will continue to use an external evaluator, you need to think about how you and your staff can develop your own evalua-tion skills. Even if you choose not to undertake major evaluations yourself, your staff team will be more effective if they have the ability to reflect on their own work, organise regular reviews and develop their critical analysis skills.

We often provide workshops in evaluation skills running in parallel with the main evaluation process and also encourage staff and management to become actively involved with the evaluation process by helping not only with the data collection but also with the interpretation of the data. We also have an aftercare policy whereby we negotiate with the client a range of possible no-charge follow-up activities of up to 5% of the price of the evaluation. These activities have included running 'debriefing' workshops on the evaluation process, providing support to staff who have to implement recommendations in the evaluation report, helping to write action plans and so on. There are also many excellent handbooks and tools for self-evaluation that we encourage project teams to use. Whilst we are not advocating that every evaluator should adopt the same policy, you should be suspicious of any evaluator who is actively encouraging a long-term dependency relationship. If your evaluator understands that a key outcome requirement is sustainability of the project past the initial funding stage, then they should be doing their bit to promote client autonomy and not simply reporting on it.

# Section 3

# Chapter 11
# Ethical considerations

As a project manager, you need to ensure that you understand what the ethical standards of your evaluator are.

As evaluation has developed as an area of study in its own right with its own theoretical bases as well as its own practice, so evaluation research has become increasingly concerned with examining – and challenging - the ethics of the evaluation intervention.

It is important for evaluators to be aware of what kinds of ethical standards they are applying and to make these explicit. As a project manager, you need to ensure you understand what the ethical standards of your evaluator are to establish that they are congruent with your own and those of your organisation so that you are not open to attack on arbitrariness.

It is obviously impossible to prescribe what the ethical stance of other individuals and evaluation organisations should be, so as an illustration of what we mean, we are including our own statement of ethics, which is shared by other members of the Capitalisation and Evaluation Research Network.

### 11.1 CERN statement of ethics.

We are committed to:

– Evaluation as an essential element in the design and planning of any project, programme or innovative process.
– Evaluation that is integral to organisational and programme activities and not 'bolted-on'.
– Evaluation that spans the whole lifecycle of a project or programme and which is formative as well as summative.
– Evaluation that is client centred, based on a non-dependency relationship and leading to long-term client autonomy and sustainability.
– Evaluation that recognises the diversity of stakeholders and responds to their different needs by offering a wide range of review and evaluation products, tools and processes.
– Evaluation as a skilled intervention and a specialist field of knowledge and practice.
– Evaluation that is ethical, transparent, professional and responsible.
– Evaluation which is informed by a range of different approaches and theoretical perspectives to ensure congruence between the review and evaluation process and the policies, processes and practices being reviewed.

Your evaluator should be able to provide a similar statement of ethics, or at least talk you through the ethical principles that underpin their work.

Ideally, your evaluator should be able to provide you with a written statement similar to the one above. However, if they cannot, don't feel shy about asking them to talk you through some of their underpinning values and ethical principles. If they are not prepared to do this – or if they have never considered the question – then maybe you should be concerned.

#### Authorship, Ownership and Censorship.

The evaluator's role will almost certainly involve the production of written documents. This may simply be recording the comments and conclusions of

others (the evaluator acting as "scribe" to a team review and evaluation activity) or it may be the critical analysis of an individual evaluator acting autonomously.

Resolve at the outset of the evaluation process, the issues of authorship, ownership, circulation and the right to censor that apply to evaluation reports and other evaluation products.

It is essential at the early stages of contracting with an external evaluator to establish the authorship and ownership of evaluation products, their circulation and the rights of reply of groups and individuals and the 'censorship' rights of the contractor. Issues of copyright and publication rights, IPR and use of purpose-designed tools also need to be resolved.

In most cases, the evaluator is a paid sub-contractor and thus the evaluation reports become the property of the contractor along with the copyright to do with as they wish. Taking an extreme example, an evaluation report may be heavily critical of a project. The project staff, having taken delivery of the report, 'own' the report and can withhold publication, edit or censor parts of it. There is little the evaluator can do in this case but can ask that their name and the name of their organisation be removed. That is, they can properly deny authorship if the text has been severely edited. What they cannot do is stop parts of the text that they have provided being used, albeit custom and practice would expect these individual parts to acknowledge authorship.

The examples given above are extreme cases and usually only result where there has been a serious breakdown in relationships. More typically, the evaluator will produce a draft report that will be circulated by the project staff for consultation before the final version is agreed.

We have also encountered disputes on ownership of evaluation materials. Our understanding is that the copyright on any questionnaires and tools designed by the evaluator for the project and specified in the contract may be the property of the contractor, whilst the Intellectual Property Rights can never be removed from the author, that is, you can always quote yourself! This is a complicated field and if it is an issue on a particular contract, specialist information should be sought in advance of the contract.

[Personally, we would have serious reservations about any evaluator who didn't make the tools they used for the evaluation freely available for the project to use in the future – and encouraged them to use them! The only exception may be if a particular tool requires a specialist level of skill to operate and interpret in which case the evaluator may be justified in worrying that it will be inappropriately used. If this is the case, they should tell you!]

The important thing is to discuss authorship of documents and ownership of evaluation products, including who has powers and rights of censorship, before they become issues. Also, it may be that problems can be avoided by having multiple evaluation products. For example, it would be inappropriate for a document in the public domain to contain explicit references to the internal dynamic of the staff team nor would such specific information be relevant to others. However, it may be useful information to feedback verbally to the people involved or as a different document.

The toolbox contains a checklist on authorship, censorship and copyright, which you may like to use as the basis of a discussion with your evaluator after which, it could be included in the contract.

### 11.2 Confidentiality

Confidentiality is often a difficult problem to resolve and you need to discuss this with your evaluator at the start.

### Anonymity

People's responses to interviews or questionnaires may vary according to the level of confidentiality they are offered – particularly if their perceptions of the project are negative or their comments are critical of the organisation or of senior management. So, you need to establish with your evaluator the level of confidentiality you feel comfortable with. At one extreme you could give an assurance that you will never want to see any of the raw data from questionnaires or interview notes nor will you ever ask for an anonymous comment to be identified. At the other end of the spectrum you could instruct the evaluator that s/he is to report back any criticism together with the name of the person that made it.

The crucial point is that the evaluator has to make the position clear to the respondents or interviewees before any exchange begins. Obviously, the level of confidentiality promised is likely to skew the outcomes of the interview and this should be taken into account. Allowing the evaluator maximum freedom is usually the best bet.

Also, it is not sufficient to assure a level of confidentiality simply by telling people that they do not have to put their names on the questionnaire. With small groups it is often easy to identify individuals by other means (nature of the response, handwriting, etc.).

Absolute confidentiality can never be assured – the most obvious reason being that if it is evident that someone is breaking the law, the evaluator has a duty to report it.

Evaluators will have to make choices, for example, a questionnaire respondent can be completely anonymous, or can be readily identified by name or identified by a code known only to a limited number of people. Who can access that code will need to be negotiated.

Sometimes, even if there is no obvious reason for identifying individuals (other than by category) there may be merit in doing so. For instance, it could be that after a large-sample questionnaire, the evaluators may want to go back to some of the respondents that have raised particular issues for a follow up interview and will therefore need some way of identifying people by name.

### Authentication

If one-to-one interviews are the main form of data collection and the opinions of the interviewees provide the bulk of the evidence on which the evaluator bases his/her judgements, then there must be some way of proving that this data is reliable and the evaluator has not made it all up. That is to say, if an evaluation report claims that *"over 50% of the policy makers interviewed felt that the project was a waste of money"* then that claim has to be justified *if required*. The established way of doing this in academic research is for the researchers to authenticate the data by transcribing the interview and then sending a copy to the interviewee for them to sign as being a true record. The authenticated transcripts are then kept so that they can be produced if necessary. The same process could be used in evaluation.

However, this can pose problems as many interviewees will say something 'off the record' but are unwilling to formally acknowledge that they have said it. Also comments made during interviews may be written in evaluation products as reported speech or as direct quotations but in both cases it may be easy to identify the authorship because of the subject matter or the particular perspective from which their observations are made.

If the evaluation is politically sensitive, or likely to be controversial, it is sensible for the evaluator to establish with the interviewee whether a remark

- Can be quoted and attributed in a report.
- Can be quoted in a report anonymously but with authenticated interview notes to back it up.
- Can be quoted anonymously and not recorded in notes.
- Can be reported rather than quoted.
- Is off-the-record and should not be quoted nor reported.

We have rather laboured this point because confidentiality is a serious issue and can cause major difficulties if handled insensitively. In reality, if there is trust between the evaluator and the project management and between the evaluator and the interviewees, then there is rarely a problem.

### Access to documents and ownership of data.

Evaluators may need to have access to project or institutional documents which have a restricted circulation and the use of material from these documents in evaluation reports should be clarified.

Ownership and confidentiality of the raw data (for example, completed questionnaires, frequency counts, interview notes etc.) can also be an issue. Evaluators will frequently want to discuss points arising from the interview notes with clients, particularly the project staff themselves, to check the significance of particular responses, to check for accuracy or to provide early feedback. Do project staff have a right to ask for the source? Whether this data is `owned' by the evaluator or the contractor should be agreed.

Who will 'own' the raw data collected by your evaluator? With your evaluator, formulate an answer to this question.

A second issue is whether either the evaluator or the project staff who provide the data can re-use it at a later date in another context and what the confidentiality limits are? For example, an evaluator recently undertook an evaluation of a project that was concerned with social exclusion issues in mainly Black and Minority Ethnic communities in Wales. One of the many data sets provided information about the levels of domestic violence. In a subsequent research report by the evaluator on domestic violence, this data was re-used and indicated that the level of domestic violence in that particular Black community was higher than in adjacent areas. The clients for the first evaluation complained because they felt that a negative use was being put to the data and argued that as it was `their' data, it should not have been re-used without permission as their report was confidential. The evaluator argued that as the report had not been mentioned, nor in fact the project, the raw data he had collected belonged to him.

These examples are fortunately not common but, nevertheless, need to be resolved.

To help you clarify the confidentiality limits before you start, we have included a Confidentiality Checklist in the toolbox

# Chapter 12
# Problems and pitfalls

Despite reading this handbook, you may still be unconvinced about the benefits of evaluation. Or you may be concerned that you will meet with resistance from others and worried about how you sell the idea to project staff or management committees. What follows are some of the major criticisms we have heard with some suggestions on how to combat them!

### "Evaluation increases the workload for project staff"

Well, yes it does! However, the increased workload should be seen as short-term investment of resources for long term gain. That is, time spent on evaluation can provide information for staff about what is working and what is not, whether they should be doing more or less of something, what is missing from the project and needs including or what needs to be abandoned. Knowing this should improve their efficiency (in terms of time and resources) and their effectiveness (in terms if achieving the outputs). Evaluation is also a way of validating success and recognising the contribution of project staff. Conversely by highlighting problems and recommending alternative courses of action, evaluation can improve future planning and reduce the likelihood of failure.

Staff may be asked to invest time in evaluation activities in various ways. Firstly, they may be asked to keep additional records or information about some aspect of their work. Secondly, staff may be involved in collecting data from beneficiaries by administering questionnaires or conducting interviews simply because they have easy access to them or have won their confidence. Thirdly, they may be asked to give up time to talk to the evaluator. Finally, they may be involved in the design of the evaluation and the processing of the data.

A negative perception of evaluation amongst project staff is common.

Some years ago we undertook some research on people's perception of evaluation. One of the findings was that the first activity listed above is the most common and the rest fall in sequence to the last, which is the least common. We have also asked staff how they perceive these activities and, unsurprisingly, there is a direct correlation between how unpopular they are and how often they are asked to undertake them! Put more explicitly, project staff resent having to keep extra records and having to do this makes them perceive evaluation negatively – yet it is the most frequent experience staff have of the evaluation process. Conversely, staff are seldom involved in the design of the evaluation process or asked to contribute to the analysis but those that have been, were positive about the experience and also perceived the evaluation process as a whole more positively.

Make clear to your staff the purpose of the evaluation, how much of their time will be needed and involve them as much as possible in the design, data analysis and feedback processes.

There are several clear messages here! If you are going to involve staff in the evaluation process then:

- Be realistic about the extra work involved and tell staff well in advance so that they can plan their time accordingly.
- Decrease the burden by incorporating evaluation activities into ongoing project activities.

– Include their input at an early stage. For example, ask them what evaluation questions would they want to ask. If you are expecting them to administer a questionnaire, discuss it with them beforehand and ask for their comments.
– Give them feedback on the data that has been collected. If they have been involved in the collection, then they are likely to be curious about the results. Ideally, they should be given a chance to look at the raw data and offer their interpretation.

.

"Money spent on evaluation could be better spent on improving the project activities."

Evaluation does improve the project activities by identifying what is working, what is not working and how the project could improve. This reduces waste and means resources can be targeted more precisely. Evaluation is part of 'good housekeeping'.

"What if we get a bad evaluation report? Why should we spend money on something that criticises what we are doing?"

Very often the answer will be because you have no choice – it may be a condition of funding. However, a less cynical answer is that by setting up an evaluation process you are making a statement to staff, funding agencies and the wider public that your project is genuine, transparent and accountable. This is in itself a good-news message that will counteract any critical comment from your evaluators.

If an evaluation identifies real problems then this should not be seen as a failure of the project but as useful information that will benefit your future work and help other projects. Also, if your agency establishes a reputation for rigorous evaluation, then your successes will also be more credible.

"We already review our work – why do we need anything more complicated?"

A good project manager will be constantly monitoring performance and making adjustments as part of day-to-day project management. Setting up a more formal system and appointing a specialist evaluator or team of evaluators, simply makes this process explicit and systematic. Depending on the skills and technical competence of the evaluator, it also provides opportunities for using more sophisticated tools and techniques that can improve the efficiency and effectiveness of existing review processes and remove some of the burden from project managers.

We are always cautious about claiming that external evaluators can be more 'objective' about a project but we are confident in saying that external evaluators have less role conflict. By this we mean that managers may be making critical judgments about their own projects but are more likely to compromise those judgments because they are inevitably influenced by other factors. For example, a manager may recognise a problem but be reluctant to identify a solution that is going to create more work for them.

Managers are also less likely to 'think outside the box' or establish causal relationships between events, especially when the cause is deeply embedded in the organisational culture of which they are part. Very often organisations have made assumptions or value judgments that have generated particular work practices that over time have come to be seen as inviolate. Somewhere

along the way the potential for change has been edited out and it is almost impossible for people within the culture to recognise this.

The remainder of this section is taken from – or based on – 'The Program Mangers Guide to Evaluation' published by the Administration on Children, Youth and Families (ACYF) of the US Department of Health and Human Services in Washington. We have found this an excellent resource and would recommend it as additional reading for project managers, especially those involved in social inclusion projects. The handbook is available free of charge on their website.

To read 'The Program Managers Guide to Evaluation' go to: www.acf.hhs.gov/programs/acyf/

### "Evaluation is just another form of monitoring."

Program managers and staff often view project evaluation as yet another way for funders to monitor projects to find out whether staff are doing what they are supposed to be doing. Project evaluation, however, is not the same as monitoring. Sometimes the information collected to monitor a program overlaps with information needed for an evaluation but the two processes ask very different questions.

There may be some overlap in activity, but Evaluation and Monitoring are two different processes.

### "Evaluation requires setting performance standards and this is too difficult."

Many project managers believe that an evaluation requires setting performance standards, such as specifying the percentage of beneficiaries who will demonstrate changes or exhibit particular behaviours. Project staff worry that if these performance standards are not met, their project will be judged a failure.

This concern is somewhat justified because often funders will require such standards to be set. However, performance standards can only be set if there is extensive base-line evaluation information on a particular programme in a variety of project settings. Without this information, performance standards are completely arbitrary and meaningless."

In our experience most funding programmes do not have sufficient information to establish these standards in any meaningful way and we would support any project manager who challenges performance standards set by funding agencies that are not based on previous research or the evidence of previous evaluations.

We would advocate "an approach to evaluation that looks at whether there has been significant change in the knowledge, attitudes, and/or behaviours of a project's participants and beneficiaries in general and whether particular characteristics of the project or the participants are more or less likely to promote change."

### Pitfalls

The final section of this chapter picks up on some of the reasons why evaluation can go horribly wrong. We start with our own experiences and follow these with pitfalls identified by the evaluation teams at ACYF and reported in their Program Managers' Handbook. We concur with all of them and make no apologies if they replicate some of the points we have covered ourselves in earlier chapters!

### Different stakeholders may have different and unresolved expectations of the evaluation process and products.

In section 5.1 we looked at the variety of people who have an interest both the process of evaluation and the evaluation products. Each stakeholder has a different perspective on the project and will want different information. It is almost impossible to please everyone within a single evaluation. You need to agree from the outset who the target group for the evaluation is or allow additional resources to cater for several different groups. Once the decision is made, this needs to be communicated to all the stakeholders.

### No clear protocols on data ownership, authorship or editorial control established in advance

Probably the worst case scenario for both an evaluator and a project manager is when the evaluator identifies major problems with the project which are directly attributable to the project manager. Whilst they can be dealt with contractually and built into the evaluation from the outset, the key to preventing these situations occurring in the first place is in the quality and nature of the relationship between the evaluator and the project manager. Time spent in building trust between you and your evaluator is a sound investment.

### Unclear lines of accountability of the evaluator

The evaluator must know to whom he or she is an answerable on a day-to-day basis and to whom they are ultimately accountable. Who is the client? Is it the project manager? The steering group or committee? The project manager's line manager or some other body? This is essential information if tensions develop between the project manager and the evaluator, for example around the extreme scenario outlined above, or if there is no agreement on the audience for the evaluation products or there is to be any shift in focus of the evaluation in the light of results obtained.

### Untimely reporting

Consider a scenario in which the evaluators submit a report six weeks after the end of the project. The report highlights major areas of concern. The evaluator had visited the project on a number of occasions and had noticed significant problems as early as the first couple of months. Further problems were identified at later stages but at no time did these surface until the submission of the formal final report. The contract did not specify an interim report. Was the evaluator behaving ethically in adopting this non-interventionist approach? Legally s/he was within their rights and certainly the situation could have been avoided if the original specification had provided for some sort of on-going feedback process. Nevertheless, we believe, the evaluator's lack of action is unprofessional. If evaluation does not contribute to the success of a project there is little reason for doing it and as such the evaluator shares that responsibility over and above any contractual arrangements. Any issues threatening the successful completion of the project should be reported as soon as identified.

### Under resourcing and/or mismanagement of evaluation time

If you are using an external evaluator remember that their contract is usually based on an agreed number of days tied to specific activities. If you expect your evaluator to attend, for example, all team meetings or all steering group

meetings then you will quickly 'use up' the allocated days. Another drain on resources is over reporting – many project managers ask for monthly or quarterly 'progress reports' from the evaluator over and above the contract requirements in terms of evaluation products. Your evaluator will cost these activities against the evaluation budget so agree the number and nature of progress reports before you start and keep these to a minimum.

Keep in mind your external evaluation budget if asking your evaluator to work outside the contractual arrangements.

### No mechanisms to ensure evaluation findings impact on future practice.

As a project manager, it is your responsibility to ensure that the lessons learned from the evaluation process or the recommendations for action are acted on – insofar as your sphere of influence extends. There is an argument which says it is of no concern to the evaluator what happens once the final evaluation product is delivered and the contract signed off. However, most professional evaluators will be happy to work with you on these issues and may have some useful ideas about how to put their suggestions into practice.

### Evaluation process may aggravate an already tense situation.

A well thought out project evaluation system in a mature and stable organisation won't create problems. However a bad evaluation experience will almost certainly have a negative effect and even a good one can aggravate an already tense situation. In an organisation plagued with instability, poorly motivated staff, ineffective management, external threats to funding or even redundancies, then evaluation is probably the last thing you need. Evaluation can definitely put a strain on already strained relationships. Unfortunately, a lot of organisations see this as the perfect time to employ an evaluator believing that s/he may tell them what's wrong and what they should do about it.

This is wishful thinking! If you really think you need outside help to cope with a crisis or deal with long standing organisational problems, you are better off employing a management consultant. Also, evaluators often get used to provide 'evidence' to justify unpalatable management decisions or used as a pawn between warring factions in a project. This is unethical and a misuse of evaluation. What evaluation can do is provide 'neutral' information which can inform management decisions that may prove to be unpopular or can contribute to reconciling opposing points of view simply by helping people understand the issues from a range of perspectives. The dividing line is very fine – don't push your evaluator into crossing it!

Don't expect your evaluator to do the job of a management consultant.

### Evaluator colluding with project manager and/or staff.

There are many roles the evaluator can occupy, depending on his or her preferred style of working, the theoretical perspective they may be basing their work on and the clients brief. These may include critical friend, observer, facilitator and so on. Similarly, teaching, coaching, mentoring, advocating and mediating are not necessarily part of the evaluator's role but are nevertheless activities evaluators may legitimately undertake during an evaluation process. The danger is that these roles may become contaminated and the activities a substitute for evaluation rather than part of it.

We have consistently stressed the importance of building trust between the project manager and the evaluator and have emphasised that evaluators have a responsibility to raise awareness of problems as they identify them. This can result in a very close and productive working relationship in which

the project manager uses the evaluator as a sounding board for possible solutions to problems the evaluator has identified or for more general advice. There are dangers, however. Firstly, staff may see the evaluator as someone 'carrying tales' to the project manager which can result in reluctance to provide the evaluator with information. Secondly, the evaluator can become over protective of the project manager and act as their 'advocate' in meetings with others. Thirdly, the evaluator can easily become involved in the internal politics of the organisation, which prevents them from being dispassionate in their judgements.
[What follows is drawn from – or based on – the ACYF Program Manager's Guide to Evaluation mentioned above.]

### Too little investment in planning.

Invest both time and effort in deciding what you want to learn from your evaluation. This is the single most important step you will take in this process. Consider what you would like to discover about your project and its impact on participants and use this information to guide your evaluation planning.

### No integration of the evaluation into ongoing activities of the project – evaluation is a bolt on.

Project managers often view evaluation as something that an outsider "does to" a project after it is over or as an activity "tacked on" merely to please funders. Unfortunately, many projects are evaluated in this way. This approach greatly limits the benefits that project managers and staff can gain from an evaluation. Planning the evaluation should begin at the same time as planning the project so that you can use evaluation feedback to inform project operations.

### Project managers pay lip service to evaluation.

Participate in the evaluation and show project staff that you think it is important. An evaluation needs the participation of the project manager to succeed. Even if an outside evaluator is hired to conduct the evaluation, project managers must be full partners in the evaluation process. An outside evaluator cannot do it alone. You must teach the evaluator about your project, your participants, your methods and your objectives. Also, staff will value the evaluation if you, the project manager, value it yourself. Talk about it with staff individually and in meetings. If you hire an outside evaluator to conduct the evaluation, be sure that this individual attends staff meetings and gives presentations on the status of the evaluation. Your involvement will encourage a sense of ownership and responsibility for the evaluation among all project staff.

### Lack of ownership of the evaluation process.

Involve as many of the project staff as much as possible and as early as possible. Project staff have a considerable stake in the success of the evaluation and involving them early on in the process will enhance the evaluation's effectiveness. Staff will have questions and issues that the evaluation can address and are usually pleased when the evaluation validates their own hunches about what does and does not work in the project. Because of their experiences and expertise, project staff can ensure that the evaluation questions, design and

methodology are appropriate for the project's beneficiaries (if they are to be surveyed) and realisable in the workplace. Furthermore, early involvement of staff will promote their willingness to participate in data collection and other evaluation-related tasks.

The participation of your staff will also be key to the effectiveness of the evaluation.

### Over ambitious evaluation.

Be realistic about the burden on yourself and your staff. Evaluations are work. Even if your evaluation calls for an outside evaluator to do most of the data collection, it still takes time to arrange for the evaluator to have access to records, administer questionnaires or conduct interviews. It is common for both agencies and evaluators to underestimate how much additional effort this involves. When project managers and staff brainstorm about all of the questions they want answered, they often produce a very long list. This process can result in an evaluation that is too complicated. Focus on the key questions that assess your project's general effectiveness or pick some very specific topics that you are concerned about.

### Insensitivity to ethical and cultural issues in an evaluation.

When you are evaluating a project that provides services or training, you must always consider your responsibilities to the beneficiaries and the community. You must ensure that the evaluation is relevant to and respectful of the cultural backgrounds and individuality of participants. Evaluation instruments and methods of data collection must be culturally sensitive and appropriate for your participants. Participants must be informed that they are taking part in an evaluation and that they have the right to refuse to participate in this activity without jeopardizing their participation in the project. Finally, you must ensure that confidentiality of participant information will be maintained as far as possible and that the level of confidentiality is communicated to all those involved.

Ensure that your evaluation - the methods of collecting data and reporting - are sensitive and respectful to your client group and local community.

# Chapter 13
# Models and theories in evaluation

As a project manager, you can manage perfectly well without this chapter as you are more likely to be concerned with the practicalities of evaluation. However, we felt it was important to include it as it will:

- Help you understand where you evaluator is coming from and the ideas that will impact on their practice.
- Help you talk to your evaluator by giving you some shared understandings.
- Help you select the type of evaluator you want.
- Demonstrate to your evaluator that you are interested in evaluation and serious about it.
- Provide a foundation for further reading.

## 13.1 Evolution of evaluation theory

Evaluation as a formal activity that we would recognise, has existed for a surprisingly long time. One of the earliest recorded was the evaluation of the effectiveness of lime-juice in preventing scurvy in sailors – commissioned by the British navy in the 18th century! The French make even earlier claim and say that the Norman armies conducted an evaluation of the relative effectiveness of the crossbow and the longbow. Unfortunately, on the basis of the evaluation findings, the management decision was to go for the crossbow and the rest, as they say, is history!

However, evaluation has only become a recognised area of academic study since about the 1960's. It is probably true to say that evaluation started as a field of practice and the theory was derived from it. As it evolved, so ideological disputes developed alongside disagreements on definitions, terminology, ethics and so on. FitzPatrick, Sanders and Worthern in 2004 identified nearly 60 different models in the 30 years between 1960 and 1990 alone. This proliferation of models was bewildering for the practitioner, especially as many of these models and the tools they generated had no obvious theoretical perspective.

Why is this a problem? Why should practitioners need a theoretical framework? Simply, a 'good' theory will set out the assumptions that it is making and on which its logic is predicated. Different theories make different assumption and generate models that will be based on different pre-conceptions and definitions of evaluation, which in turn lead to very different practices.

### Deriving a taxonomy of evaluation approaches.

Many researchers have tried to make sense of this huge diversity of models and theories and to find some way of classifying them. However, even they could not agree so now we have the problem of trying to classify the classification systems!

All this is by way of saying that what follows is only one framework for distinguishing between different theories of evaluation and you may well come across others. This framework, which we find comprehensive and use-

able, was devised by FitzPatrick, Sanders and Worthen (2004) who we have quoted at length.

### Philosophical /ideological differences.

Approaches to evaluation may differ fundamentally because their underpinning philosophy or ideological base is different. That is, different evaluation theories will be based on different assumptions about the way the world works and so the models and practices based on those theories will be different as well. By and large, we can locate them on a continuum from objectivist to subjectivist.

Objectivism is equivalent to the empirical tradition in scientific research (positivism) and focuses on data collection and analysis techniques that produce results that are reproducible and verifiable by other evaluators and to generate conclusions that are evidence based and which can be 'scientifically' justified. So the evaluation is 'external' to the evaluator who is simply someone technically competent and proficient in the application of procedures. Subjectivism is based on:

> "...an appeal to experience rather than to scientific method. Knowledge [of the evaluator] is conceived as being largely tacit rather than scientific."
> (House 1980 in FitzPatrick, Sanders and Worthen 2004)

The validity of a subjectivist evaluation depends on the relevance of the evaluators' background, their experience and expertise, the keenness of their perceptions and their insightfulness in generating interpretations and conclusions. Thus, the evaluation procedures are 'internal' to each evaluator and are not explicitly understood or reproducible by others.

Until 20 years ago, objectivism in evaluation was a goal to be aspired to. However, the same criticisms levelled at the usefulness of positivism in the social sciences in general were also applied to objectivism in evaluation.

Campbell (1984) summed it up:

> "twenty years ago positivism dominated the philosophy of science...Today the tide has completely turned among the theorists of science in philosophy, sociology, and elsewhere. Logical positivism is almost universally rejected."

This point of view has been upheld by many writers on evaluation and even if it is not universally subscribed to, probably represents a general trend. The major argument is that unlike traditional scientific research, evaluation has to deal with complex phenomena in real world settings, take into account a multiplicity of stakeholders, unstable and unpredictable systems and requires a high level of human interactivity.

The other criticism is that objectivism depends for its validity on its 'scientific' methodology and is only credible if you happen to value that methodology. We would argue that objectivism conceals hidden values and biases of which many evaluators are unaware – even the choice of data collection techniques and instruments is not value-neutral but this is not recognised or else ignored by many so-called objective evaluations.

Despite the reaction of the theorists, however, the message does not seem to have filtered through to the client base and the majority of evaluation con-

sumers, particularly in education (and the public sector in general), are still wedded to the idea of objective evaluation and 'finding the facts'.

The major criticism is that subjectivist evaluation often leads to contradictory conclusions that cannot be reconciled because the processes which led to the conclusions is largely inside the evaluators head and so cannot be replicated.

### Differences based on defining value or worth.

We can also distinguish between different theoretical approaches depending on how they define value and make judgements, rather than on their philosophical differences. This time the continuum extends from 'utilitarian' to 'intuitionist-pluralist'.

'Utilitarianism' is a philosophy based on maximising happiness in society. Utilitarian approaches to evaluation are based on the premise that the best programmes are those that achieve the greatest good for the greatest number. The evaluator will try and assess overall impact in terms of total group gains by using average outcome scores against the criteria selected to determine worth. Again, governments and the public sector tend to be adherents of this type of evaluation as it lends itself to large-scale comparisons of programmes and mass aggregation of data. Managers and public programme administrators tend to be the main audiences.

According to FitzPatrick et al, the intuitionist-pluralist approach is at the other end of the spectrum and is based on the premise that value depends on the impact of a programme on each individual and the 'greatest good' is that which maximises the benefits for all stakeholders. "This evaluation focus will be on the *distribution* of *gains* (for example by cultural or sub-cultural demographic groups such as ethnicity or gender or age) or distribution of benefit across stakeholders (e.g. learners, administrators, delivery agencies, funding bodies, the public). There can be no common index of "good" but rather a plurality of criteria and judges. The evaluator is no longer an impartial 'averager' but a portrayer of different values and needs. The merit or worth of any programme depends on the values and perspectives of whoever is judging it and all stakeholders are legitimate judges."

### Methodological differences.

Don't assume that there is a fundamental divide between qualitative and quantitative approaches to evaluation.

Although there is a strong correlation between an evaluator's ideological approach and the methodology and techniques they will use (because of necessity one drives the other), there are other major divides based on methodological differences that are not necessarily rooted in a particular philosophical approach. For example, many evaluators (both theoreticians and practitioners) and also many clients tend to view qualitative and quantitative approaches as different paradigms. We do not subscribe to this view, believing that this is not a fundamental divide but simply a way of describing evaluation approaches by types of data that are used. Nevertheless, we recognise this as an important distinction for others and one that impacts on the overall evaluation methodology and the tools used.

### Differences according to discipline or field of application.

Evaluation is a relatively young field and still draws heavily on methodologies adapted from anthropology, sociology, psychology, philosophy, economics and mathematics. One of the consequences is that evaluation approaches

can be grouped around their parent discipline so we tend to find 'mathematical approaches' or 'sociological approaches'. More recently the search for new models has widened its net and evaluation theorists such as Smith (1981) are trawling newer disciplines such as investigative journalism, photography, storytellling, philosophical analysis, forensic pathology and literary criticism for new ideas.

Evaluation theory has also developed in a social context and practitioners work in different cultures, different sectors, with different target groups and different audiences. Consequently, different approaches and models have tended to emerge based on these factors. For example, 'education programme' evaluation has developed along a different trajectory than, for example, the health services. Whilst many writers would argue that this is not a true theoretical divide, 'theory-in-practice' is a powerful determinant of evaluation approach and also stakeholders perceptions and expectations of the evaluation process.

### Differences in practice.

The above distinctions are all based (loosely) on theoretical divisions. However, FitzPatrick et al also point out that differences in approach can be practice-driven.

- Firstly, evaluators disagree about whether evaluators should simply provide information so that decision makers can make the value judgements. Others, would say that the evaluator's report to decision makers is incomplete if it does not contain a value judgement.
- Secondly, evaluators differ in their perception of their own role and their place in the evaluation process. Who has authority and responsibility for the evaluation and to whom should the evaluator be accountable and answerable? If one evaluator sees his role as a 'critical friend' and another as 'inspector' or 'judge', then this will obviously influence the way they conduct an evaluation and also the conclusions they draw.
- Thirdly, evaluators will be limited by their prior experience both in evaluation and also by their own discipline or professional background. Evaluation skills and knowledge are cumulative. Previous exposure to frequently recurring problems will affect the way an evaluator works. On the one hand it will probably mean the evaluator is quicker to detect problems, to identify issues of concern and make more insightful judgements. On the other hand, it will also mean that the evaluator's perceptions in a new situation are unlikely to be 'neutral'.
- Fourthly, evaluators have different views about what skills and expertise they should possess. Evaluators are frequently chosen on the basis of their expertise or practitioner base in the field being evaluated rather than on the basis of their skills and experience as an evaluator. This is gradually changing but as evaluation is becoming increasingly professionalised and recognised as a specialist area in its own right, so professional evaluators are becoming specialised within the area. Some evaluators would argue that specialist knowledge of the field being evaluated is a pre-requisite for the credibility of the whole process of evaluation. Others claim that not only is this unnecessary but can, on occasions, be unhelpful.

### A classification system

The above analysis is interesting and helps understand the major theoretical divides in evaluation. However, it does not get us far in terms of systematically examining the variation between particular evaluation approaches because although those approaches could be positioned on each of the above 'dimensions', their location would vary from one dimension to another. The next section tries to provide some answers.

Many evaluation theorists have attempted this but we are going to stick with the solution put forward by Fitzpatrick, Sanders and Worthen (1983). We are proposing to use their work – with some modifications - partly in the interests of consistency (having referenced them heavily so far) and partly because they set out very clearly the thinking and rationale underpinning their classification system.

For the purist it is a less than satisfactory taxonomy as the approaches do not necessarily differ from one another along the same dimension. However, they are pragmatic as they conveniently represent the major clusters of models and approaches in use today.

### A Classification Schemata for Evaluation Approaches

FitzPatrick et al identify 5 major clusters of evaluation approaches:

- Objectives oriented approaches
- Management oriented approaches
- Consumer oriented approaches
- Expertise oriented approaches
- Participant oriented approaches

However, to this we propose to add Van der Knapp's 'learning oriented approach'

These 6 categories fall more or less along a continuum from utilitarian to intuitionist-pluralist so there is some logical basis in addition to its convenience and accessibility.

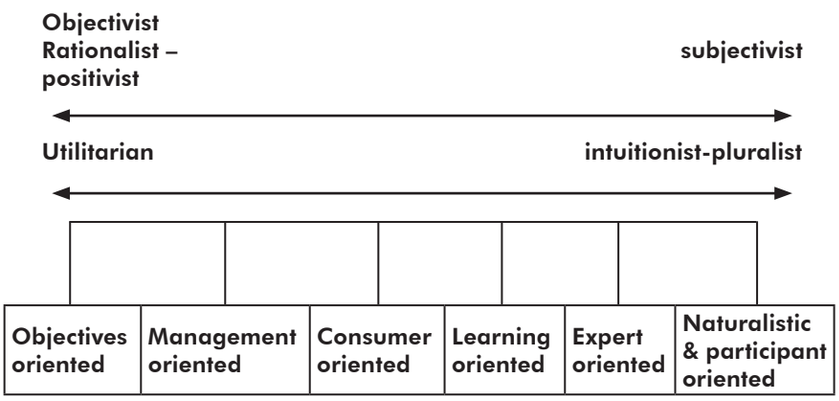


*Figure 5: Classification schemata for evaluation approaches*

### Objectives orientated evaluation approaches

Objectives-orientated evaluation is based on the idea that the purposes, goals or targets of a project are determined at the start and the evaluation process should establish whether these have actually been achieved – and, if not, why not. It is very similar to another approach known as 'a systems approach'

to evaluation and both are very popular with public sector agencies who are concerned with justifying expenditure and performance measurement. It is sometimes called 'goal-driven' evaluation, in contrast with other approaches, which are called 'goal-free'.

There are many examples of objectives orientated models; the earliest is probably Tyler's and more recently, Provus's Discrepancy Model.

The disadvantages are that this sort of approach can miss important outcomes if they were not included in the original objectives nor does it challenge the value of the objectives themselves

### Management orientated evaluation approaches.

The management-orientated approach to evaluation is meant to serve decision makers. Its rationale is that evaluation information is an essential part of good decision making and that the evaluator can be most effective by focussing the evaluation products on the needs of managers, policymakers, administrators and practitioners.

A management-orientated evaluation of your project should inform and assist management decision making.

Developers of this approach have traditionally relied on a systems approach to evaluation in which decisions are made about inputs, processes and outputs based on logic models and cybernetic theory. However, more recent developments have highlighted different levels of decision and decision makers and have focussed on who will use the evaluation results, how they will use them and what aspect(s) of the system they are making decisions about.

Not surprisingly, it is the model preferred by many managers and management committees but the downside is that the needs of other stakeholders are ignored.

Stufflebeam's CIPP model is one of the most popular in management-orientated evaluation.

### Consumer orientated approaches.

Consumer orientated approaches to evaluation adopt the perspective of the end user of whatever service or product is being provided. For this reason they tend to be summative, rather than formative and are concerned primarily with product evaluation. Consumer-orientated evaluation relies heavily on criteria referenced evaluation techniques such as benchmarking or kite marking and is understandably popular with standards agencies and 'watchdog' organisations. Michael Scrivens 'Key Evaluation Checklist' is probably the best-known example.

The major disadvantage of a consumer-orientated approach is that it is a 'backward-mapping' approach and does not help make predictions about future impacts. It also tends to play down the nature of human interaction with the products being evaluated.

### Expertise orientated approaches

Expertise orientated evaluation is based on the notion of "connoisseurship" and criticism and relies on the subjective professional judgement and expert knowledge of the evaluator. This is the oldest form of evaluation and is still very popular despite its limitations.

An expertise-orientated evaluation an evaluation that relies heavily on the expertise and subjective judgement of your evaluator.

Expertise-orientated evaluation may be formal or informal, based on individual expertise or, more usually, on the collective expertise of a panel. The opinions of multiple experts is popularly believed to minimise bias, though this does not always follow! It relies far less on external tools and instruments

than other forms of evaluation and more on the experience and wisdom of the evaluator.

Many public systems are based on expertise orientated evaluation – for example the jury system, school inspection system, licensing agencies, review boards, the refereeing system for academic journals, national commissions and enquiries and so on.

Many organisations expect this type of evaluation if they employ an external evaluator and the notion of evaluation by 'peer review' is still the dominant model in most professional associations.

The disadvantages are obviously the high reliance on the assumed expertise of the evaluator and a lack of explicit and published standards. Also, the credibility of results is affected by the status of the evaluator but equally the credibility of the evaluator is often affected by the results.

### Learning-orientated evaluation approaches.

This is a relatively new group of approaches and not one that was included in FitzPatrick et al's classification. Nevertheless we have included it because it is an approach that we personally use more than any other.

The operating principle is that the purpose of evaluation is to contribute to some form of collective or organisational learning. Different models within this overall approach are based on different theories and types of learning including 'corrective' or behavioural learning, cognitive learning and social learning. The outputs and processes of the evaluation form the inputs of the learning.

The pioneer of work in this field was Peter Van der Knaap. More recently we have extended the approach to include evaluation as a contributor to knowledge creation in an organisation.

Learning-orientated evaluation approaches are still not widespread but are beginning to gather momentum in the social agency sector, in education establishments and in voluntary organisations.

The main limitations of this approach is that it does not lend itself to "mass surveys" as it relies heavily on personal interaction between the evaluator and the project team and the evaluator's understanding of the learning needs of the organisation. Also, within this overall approach there are very disparate models, some requiring a high level of commitment to the process, which may be lacking.

### Participant-orientated evaluation approaches.

"bureaucrats tend to hate it because of its lack of 'objectivity' and because the outputs of the evaluation are unpredictable".

An increasingly popular approach that differs fundamentally from all the others as it takes the needs of project participants as its starting point. This is not to say that the other approaches ignore participant needs but that for the most part benefits for participants represent the end point of the evaluation and not the beginning.

Participants are not simply the direct beneficiary target group of a project but will also include other stakeholders and *potential beneficiaries*. Thus, an educational project for women returners would include the learners themselves, the project staff, the management team and the funders but may also include the wider community, the learners families, the schools attended by the learners' children, childcare agencies or whatever.

Participant-orientated evaluation does not usually follow a formal plan drawn up in advance; rather it looks for patterns in the data as the evaluation

progresses. Data is gathered in a variety of ways, using a range of techniques and culled from many different sources. Understandings grow from observation and bottom up investigation rather than rational deductive processes. The evaluator's role is to represent multiple realities and values rather than singular perspectives.

Participant-orientated evaluation includes many sub-groups that share all or some of the above characteristics including Responsive Evaluation, Naturalistic Evaluation, Utilisation Focussed evaluation and Empowerment Evaluation. Of all the models, probably the best known and one of the most useful is Stakes Countenance Framework.

Criticisms of this approach are many; bureaucrats tend to hate it because of its lack of 'objectivity' and because the outputs of the evaluation are unpredictable. It is difficult to cost and control. Without a very experienced evaluator to facilitate the process, it can degenerate from an 'organic' approach to one which is chaotic and lacking in focus. Also, there may be concentration on process at the expense of outputs.

# Chapter 14
# And finally … Evaluating your evaluation system

This final chapter will help you decide if your evaluation is right for your project.

By now you should be in a position to plan, organise and manage an evaluation process and work alongside an evaluator or an evaluation team. But the final question is: how do you know that your evaluation system measures up? This handbook would not be complete without a final section on evaluating your evaluation systems.

Meta-evaluation (evaluation evaluation) is a complex area and outside the scope of this book. However, we can give you one tool, which might help you to look critically at your evaluation system.

Imagine you were in a shop that sold evaluation systems 'off-the-peg' – what would criteria would you use if you went in to buy one?

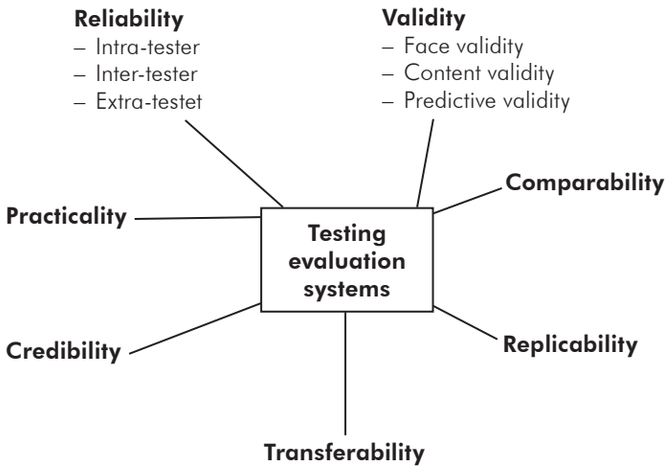


*Figure 6: Testing evaluation systems*

### Reliability

Reliability is a measure of consistency. A robust evaluation should be reliable, that is, it should yield the same results irrespective of who is conducting it or the environmental conditions under which it is taking place. Intra-tester reliability simply means that if the same evaluator is evaluating your project his or her judgement should be consistent and not influenced by, for example, another project they might have just visited or whether they feel unwell or just in a bad mood! Inter-tester reliability means that if two different evaluators were given exactly the same questions, data collection tools, output data and so on, their conclusions should also be the same. Extra-tester reliability means that the evaluator's conclusions should not be influenced by extraneous circumstances, which should have no bearing on the evaluation object.

### Validity

Validity is a measure of 'appropriateness' or 'fitness for purpose'. There are three sorts of validity. Face validity implies a match between *what* is being evaluated or tested and *how* that is being done. For example, if you are evalu-

ating how well someone can bake a cake or drive a car, then you would probably want them to actually do it rather than write an essay about it! Content validity means that *what* you are testing is actually relevant, meaningful and appropriate and there is a match between what the project is setting out to do and what is being evaluated. If an evaluation system has predictive validity it means that the results are still likely to hold true even under conditions that are different from the test conditions. For example, performance evaluation of airline pilots who are trained to cope with emergency situations on a simulator must be very high on predictive validity.

### Replicability

Ideally an evaluation should be carried out and documented in a way which is transparent and which allows the evaluation to be replicated by others to achieve the same outcomes. Some 'subjectivist' approaches to evaluation would disagree, however.

### Transferability

Although each evaluation should be designed around a particular project, a good evaluation system is one that could be adapted for similar projects or could be extended easily to new activities of your project. That is, if your project evolves and changes over a period of time in response to need, it would be useful if you didn't have to rethink your entire evaluation system. Transferability is about the shelf-life of the evaluation and also about maximising its usefulness.

### Credibility

People actually have to believe in your evaluation! It needs to be authentic, honest, transparent and ethical. If you have even one group of stakeholders questioning the rigour of the evaluation process or doubting the results or challenging the validity of the conclusions, the evaluation loses credibility and is not worth doing.

### Practicality

This means simply that however sophisticated and technically sound the evaluation is, if it takes too much of people's time or costs too much or is cumbersome to use or the products are inappropriate then it is not a good evaluation!

### Comparability

Although this handbook has been written to help you design your own evaluation system which will be tailored to meet the requirements of your particular project, a good evaluation system should also take into account the wider 'evaluation environment' in which your project is located. For example, if you are working for a government department where objective-based evaluation, quantitative methods and the use of numeric performance indicators is the norm, then if you undertake a radically different type of evaluation you may find that your audience will be less receptive and your results less acceptable. Similarly, if your project is part of a wider programme and all the other projects are using a different system then this could mean that your input is ignored simply because it is too difficult to integrate.

Also, if you are trying to compare your project's performance from one year to the next or compare your project with another, then this needs to be taken into account.

# Chapter 15
# Toolkit

## Sample Outline for Evaluation Plan

Adapted the Program Managers Guide to Evaluation published by US Department of Health and Human Services, Administration for Children and Families.

So that both you and your evaluator have a clear idea of what the evaluation will look like and so that you can monitor it, you will need a written evaluation plan. Ideally, this should be prepared at the main project planning stage. However, in practice it is likely to be prepared jointly by you and your evaluator after he or she is appointed – which is almost always after the project has started. One way of preparing a plan is to use the on-line Evaluation Mentor (see p. 36) An alternative is to use the following outline plan

I. Evaluation framework

A. What you are going to evaluate.
   1. Program model (assumptions about target population, interventions, immediate outcomes, intermediate outcomes, and final outcomes).
   2. Program implementation objectives (stated in general and then measurable terms).
      a. What you plan to do and how
      b. Who will do it
      c. Participant population and recruitment strategies
   3. Participant outcome objectives (stated in general and then measurable terms).
   4. Context for the evaluation.

B. Questions to be addressed in the evaluation.
   1. Are implementation objectives being attained? If not, why (that is, what barriers or problems have been encountered)? What kinds of things facilitated implementation?
   2. Are participant outcome objectives being attained? If not, why (that is, what barriers or problems have been encountered)? What kinds of things facilitated attainment of participant outcomes?
      a. Do participant outcomes vary as a function of program features? (That is, which aspects of the program are most predictive of expected outcomes?)
      b. Do participant outcomes vary as a function of characteristics of the participants or staff?

C. Timeframe for the evaluation.
   1. When data collection will begin and end
   2. How and why timeframe was selected

II. Evaluating implementation objectives – procedures and methods.

A. Evaluation design
B. Objective 1 (state objective in measurable terms)

1. Type of information needed to determine if objective 1 is being attained and to assess barriers and facilitators.
2. Sources of information (that is, where you plan to get the information including staff, participants, program documents). Be sure to include your plans for maintaining confidentiality of the information obtained during the evaluation.
3. How sources of information were selected.
4. Time frame for collecting information.
5. Methods for collecting the information (such as interviews, paper and pencil instruments, observations, records reviews).
6. Methods for analysing the information to determine whether the objective was attained (that is, tabulation of frequencies, assessment of relationships between or among variables).

C. Repeat this information for each implementation objective being assessed in the evaluation.

III. Evaluating participant outcome objectives – procedures and methods.

A. Evaluation design
B. Objective 1 (state outcome objective in measurable terms)
1. Types of information needed to determine if objective 1 is being attained (that is, what evidence will you use to demonstrate the change?)
2. Methods of collecting that information (for example, questionnaires, observations, surveys, interviews) and plans for pilot-testing information collection methods.
3. Sources of information (such as program staff, participants, agency staff, program managers, etc.) and sampling plan, if relevant.
4. Timeframe for collecting information.
5. Methods for analyzing the information to determine whether the objective was attained (i.e., tabulation of frequencies, assessment of relationships between or among variables using statistical tests).

C. Repeat this information for each participant outcome objective being assessed in the evaluation.

IV. Procedures for managing and monitoring the evaluation.

A. Procedures for training staff to collect evaluation-related information.
B. Procedures for conducting quality control checks of the information collection process.
C. Timelines for collecting, analysing, and reporting information, including procedures for providing evaluation-related feedback to program managers and staff.

www.acf.dhhs.gov/programs/core/pubs_reports/chapter

## Who needs to be included in the evaluation process?

| Provide Information | Receive information | |
|---|---|---|
| | | **Paymasters** |
| | | External funding agencies |
| | | Technical assistance or other intermediaries or gatekeepers e.g. auditors, national co-ordinating units, monitoring committees, other – please specify |
| | | Other institutional departments |
| | | **Promoters** |
| | | Governing bodies |
| | | Executive boards |
| | | Institutional senior managers, middle managers or other departmental staff |
| | | Partners |
| | | Other (please specify) |
| | | **Participants** |
| | | Beneficiaries, trainees, individual end users |
| | | Community groups/user groups |
| | | Other agencies (please specify) |
| | | **Policy makers** |
| | | Transnational |
| | | National |
| | | Regional |
| | | Local |
| | | Institutional |
| | | Other |
| | | **Providers** |
| | | Project staff |
| | | Steering groups / management committees |
| | | External training providers |
| | | Other sub-contractors |
| | | **Practitioners** |
| | | Working in the same field / subject |
| | | Working with the same clients |
| | | Researchers |

**Secondary data checklist**

Please indicate whether you have or can obtain access to the following documents.

| Project | Evaluator | |
|---|---|---|
| ☐ | ☐ | Local demographic data |
| ☐ | ☐ | Local structural plans |
| ☐ | ☐ | Local strategy plans |
| ☐ | ☐ | Socio-economic profile of area |
| ☐ | ☐ | Other statistical data |
| ☐ | ☐ | Institutional profile |
| ☐ | ☐ | Institutional and/or departmental policy and strategy documents |
| ☐ | ☐ | Background literature related to area of activity including journal articles, conference reports, unpublished research reports, working papers etc. |
| ☐ | ☐ | Relevant European, national or regional policy and strategy documents |
| ☐ | ☐ | Reports or case studies of projects undertaking similar work |
| ☐ | ☐ | Project application |
| ☐ | ☐ | Internal project reports |
| ☐ | ☐ | Minutes of steering group meetings, staff meetings, management meetings etc. |
| ☐ | ☐ | Feasibility studies and pilot studies |
| ☐ | ☐ | Previous evaluation reports for similar projects |
| ☐ | ☐ | Maps of area |

## Basics of Developing Questionnaires

Adapted from Carter McNamara, PhD 1999 written for the Management Assistance Program for Nonprofits.

### Key Preparation.

Before you start to design your questions, articulate clearly what problem or need will be addressed using the information to be gathered by the questions. Review why you are doing the evaluation and what you hope to accomplish by it. This provides focus on what information you need and, ultimately, on what questions should be used.

### Directions to Respondents.

- Include a brief explanation of the purpose of the questionnaire.
- Include clear explanation of how to complete the questionnaire.
- Include directions about where to provide the completed questionnaire.
- Note conditions of confidentiality, e.g., who will have access to the information, whether you are going to attempt to keep their answers private and only accessed by yourself and/or someone who will collate answers.

(Remember you cannot guarantee confidentiality. If a court of law asked to see the answers, you would not be able to stop them. However, you can assure people that you will make every reasonable attempt to protect access to their answers. You should consider using an informed consent form, as well.)

### Content of Questions.

- Ask about what you need to know, that is, get information in regard to the goals or ultimate questions you want to address by the evaluation.
- Will the respondent be able to answer your question, that is, do they know the answer?
- Will respondents want to answer the question, that is, is it too private or silly?

### Wording of Questions.

- Will the respondent understand the wording, that is, are you using any slang, culturally-specific or technical words?
- Are any words so strong that they might influence the respondent to answer a certain way? Attempt to avoid use of strong adjectives with nouns in the questions, for example, 'highly effective government,' 'prompt and reliable,' etc.
- To ensure you're asking one question at a time, avoid use of the word 'and' in your question.
- Avoid using 'not' in your questions if you're having respondents answer 'yes' or 'no' to a question. Use of 'not' can lead to double negatives and cause confusion.
- If you use multiple-choice questions, be sure your choices are mutually exclusive and encompass the total range of answers. Respondents should not be confused about whether two or more choices appear to mean the same thing. Respondents should also not be in the position of having a clearly preferred answer that is missing from the list of possible answers to the question.

Order of Questions.

- Be careful not to include so many questions that potential respondents are dissuaded from responding.
- Try to increase the respondents' motivation to complete the questionnaire. Start with fact-based questions and then go on to opinion-based questions, for example, ask people for demographic information about themselves and then go on to questions about their opinions and perspectives. This gets respondents engaged in the questionnaire and warmed up before more challenging and reflective questions about their opinions.
  (If they can complete the questionnaire anonymously, indicate this on the form where you ask for their name.)
- Attempt to get respondents' commentary in addition to their ratings, for example, if the questionnaire asks respondents to choose an answer by circling / ticking or provide a rating, ask them to provide commentary that explains their choices.
- Include a question to get respondents' impressions of the questionnaire itself. For example, ask them if the questionnaire was straightforward to complete („yes" or „no), and if not, to provide suggestions about how to improve the questionnaire.
- Pilot or test your questionnaire on a small group of clients or fellow staff. Ask them if the form and questions seemed straightforward. Carefully review the answers on the questionnaires. Does the information answer the evaluation questions or provide what you want to know about the program or its specific services? What else would you like to know?
- Finalise the questionnaire according to results of the pilot. Put a date on the form so you can keep track of all future versions.

2233 University Avenue West, Suite 360
St. Paul, Minnesota 55114 (651) 647-1216


## Conducting Focus Groups

Adapted from Carter McNamara, PhD 1999 written for the Management Assistance Program for Nonprofits

Focus groups are a powerful means to evaluate services or test new ideas. Basically, focus groups are interviews but of 6-10 people at the same time in the same group. You can get a great deal of information during a focus group session.

Preparing for Session.

- Identify the major objective of the meeting.
- Carefully develop five to six questions (see below).
- Plan your session (see below).
- Call potential members to invite them to the meeting. Send them a follow-up invitation with a proposed agenda, session time and list of questions the group will discuss. Plan to provide a copy of the report from the session to each member and let them know you will do this.
- About three days before the session, call each member to remind them to attend.

### Developing Questions.

- Develop five to six questions – The session should last 60-90 minutes. In this time, one can ask at most five or six questions.
- Always first ask yourself what problem or need will be addressed by the information gathered during the session, e.g., examine if a new service or idea will work, further understand how a program is failing, etc.
- Focus groups are basically multiple interviews. Therefore, many of the same guidelines for conducting focus groups are similar to conducting interviews (see the Basics of Conducting Interviews).

### Planning the Session.

- Scheduling - Plan meetings to be one to 1.5 hours long. Over lunch seems to be a very good time for some people to attend.
- Setting and Refreshments - Hold sessions in a conference room, or other setting with adequate air flow and lighting. Arrange chairs so that all members can see each other. Provide name-tags for members as well. Provide refreshments, especially if the session is held over lunch.
- Ground Rules - It‘s critical that all members participate as much as possible, which means encouraging the quiet ones but keeping the session moving along and generating useful information. Because the session is often a one-time occurrence, it‘s useful to have a few, short ground rules. Consider the following three ground rules:
  - Keep focused.
  - Maintain momentum.
  - Get closure on questions.
- Agenda - Consider the following agenda: welcome, review of agenda, review of goal of the meeting, review of ground rules, introductions, questions and answers, wrap up.
- Membership - Focus groups are usually conducted with 6-10 members who have some similar nature, e.g., similar age group, status in a project, etc. Select members who are likely to be participative and reflective. Attempt to select members who don‘t know each other.
- Plan to record the session with either an audio or audio-video recorder. Don‘t count on your memory. If this isn‘t practical, involve a co-facilitator who is there to take notes.

### Facilitating the Session.

- Major goal of facilitation is collecting useful information to meet goal of meeting.
- Introduce yourself and the co-facilitator, if used.
- Explain the means of recording the session.
- Carry out the agenda – (See „agenda“ above).
- Carefully word each question before that question is addressed by the group. Allow the group a few minutes for each member to carefully record their answers. Then, facilitate discussion around the answers to each question, one at a time.
- After each question is answered, carefully reflect back a summary of what you heard (the note taker may do this).
- Ensure even participation. If one or two people are dominating the meeting, then call on others. Consider using a round-table approach, including going in one direction around the table, giving each person a minute to

answer the question. If the domination persists, note it to the group and ask for ideas about how the participation can be increased.
– Closing the session - Tell members that they will receive a copy of the report generated from their answers, thank them for coming and adjourn the meeting.

### Immediately After Session.

– Verify if the tape recorder, if used, worked throughout the session.
– Make any notes on your written notes, for example, to clarify any scratching, ensure pages are numbered, fill out any notes that don't make sense
– Write down any observations made during the session. For example, where did the session occur and when, what was the nature of participation in the group? Were there any surprises during the session? Did the tape recorder break?

## Guidelines for conducting interviews

Adapted from Carter McNamara, MBA, PhD 1999 who in turn adapted some of the material from Michael Patton's book, „Qualitative Evaluation and Research Methods" (Sage Publications, 1990).

### Introduction

Interviews are particularly useful for getting the story behind a participant's experiences or for pursuing in-depth information around a topic. Interviews may also be a follow-up to certain questionnaire respondents who have provided 'interesting' comments to investigate their responses further. Usually open-ended questions are asked during interviews.

Before you start to design your interview questions and process, clearly articulate to yourself what problem or need is to be addressed using the information to be gathered by the interviews. This helps you keep clear focus on the intent of each question.

### Preparation for Interview

– Choose a setting with little distraction. Avoid bright lights or loud noises, ensure the interviewee is comfortable (you might ask them if they are), etc. Often, they may feel more comfortable at their own places of work or at home.
– Explain the purpose of the interview.
– Address terms of confidentiality. (Be careful. You cannot promise absolute confidentiality). Explain who will get access to their answers and how their answers will be analysed. If their comments are to be used as quotes, get their permission to do so and ask whether these quotes should be attributed or anonymous.
– Explain the format of the interview. Explain the type of interview you are conducting and its nature. If you want them to ask questions, specify if they're to do so as they have them or wait until the end of the interview.
– Indicate how long the interview usually takes.
– Tell them how to get in touch with you later if they want to.
– Ask them if they have any questions before you both get started with the interview.
– Don't count on your memory to recall their answers. Ask for permission to record the interview or bring along someone to take notes.

## Types of Interviews

### Unstructured

This is usually an informal, conversational interview – no predetermined questions are asked, in order to remain as open and adaptable as possible to the interviewee's nature and priorities. During the interview, the interviewer 'goes with the flow'.

### Semi-structured

This is usually based on an interview guide sheet and is intended to ensure that the same general areas of information are collected from each interviewee. This provides more focus than the conversational approach, but still allows a degree of freedom and adaptability in getting information from the interviewee. It is useful when comparisons have to be made but in situations where respondents may be from different cultures, different organizational positions or very different demographic groups.

### Structured.

These can be of two sorts:

Standardised, open-ended interviews in which the same open-ended questions are asked to all interviewees (an open-ended question is where respondents are free to choose how to answer the question, that is, they don't select 'yes' or 'no' or provide a numeric rating, etc.); this approach facilitates faster interviews that can be more easily analysed and compared.

Closed, fixed-response interview - where all interviewees are asked the same questions and asked to choose answers from among the same set of options. This format is useful for those not practiced in interviewing. However, you may want to think about using a questionnaire under these circumstances.

### Types of Topics in Questions

Patton notes six kinds of questions. One can ask questions about:

- Behaviours – about what a person has done or is doing.
- Opinions/values – about what a person thinks about a topic.
- Feelings – note that respondents sometimes respond with „I think ..." so be careful to note that you're looking for feelings.
- Knowledge – to get facts about a topic.
- Sensory – about what people have seen, touched, heard, tasted or smelled.
- Background/demographics – standard background questions, such as age, education, etc.

Note that the above questions can be asked in terms of past, present or future.

### Sequence of Questions

- Get the respondents involved in the interview as soon as possible.
- Before asking about controversial matters (such as feelings and conclusions), first ask about some facts. This gives respondents time to engage in the interview before warming up to more personal matters.

- Intersperse opinion based and feeling-based questions throughout the interview to avoid long lists of fact-based questions, which tend to leave respondents disengaged.
- Ask questions about the present before questions about the past or future. It's usually easier to talk about the present and then work into the past or future.
- The last questions might be to allow respondents to provide any other information they want to add and their impressions of the interview.

### Wording of Questions

- Wording should be open-ended. Respondents should be able to choose their own terms when answering questions.
- Questions should be as neutral as possible. Avoid wording that might influence answers, for example, evocative, judgmental wording.
- Questions should be asked one at a time.
- Questions should be worded clearly. This includes knowing any terms particular to the project or the respondents' culture.
- Be careful of asking 'why' questions. This type of question can imply a cause-effect relationship that may not truly exist. These questions may also cause respondents to feel defensive and that they have to justify their response, which may inhibit their responses to this and future questions.

### Carrying out the Interview

- Occasionally verify the tape recorder (if used) is working.
- Ask one question at a time – avoid multiple questions.
- Attempt to remain as neutral as possible. Show respect to the interviewee but do not collude with things they may tell you by showing an overly positive response to certain answers, a strong emotional reaction or too high a level of empathy which may encourage skewed responses.
- Encourage responses with occasional nods of the head, 'uh huhs', etc.
- Be careful about appearances when note taking. That is, if you jump to take a note, it may appear as if you're surprised or very pleased about an answer, which may influence answers to future questions.
- Provide transition between major topics, for example, "we've been talking about (some topic) and now I'd like to move on to (another topic)."
- Don't lose control of the interview. This can occur when respondents stray to another topic, take so long to answer a question that times begins to run out, or even begin asking questions to the interviewer.

### Immediately After Interview

- Verify if the tape recorder, if used, worked throughout the interview.
- Make any notes on your written notes, for example, to clarify any scratchings, ensure pages are numbered, fill out any notes that don't make senses, etc.
- Write down any observations made during the interview. For example, where did the interview occur and when, was the respondent particularly nervous at any time?
- Were there any surprises during the interview? Did the tape recorder break?

## Evaluation products

The following list gives a range of options that can be effective vehicles for making available the outputs of an evaluation to different users. The list also allows you to indicate 'By whom?' and to specify the target audience and time scale.

| **Evaluation output checklist** | | | |
|---|---|---|---|
| **Who will be involved?** | | **By when?** | **What?** |
| **Evaluator** | **Other** | **Delivery Date** | |
| | | | Written reports (individual or in combination) |
| | | | Position audit (before and after) |
| | | | Final report |
| | | | Executive summary |
| | | | SWOT analysis |
| | | | Research paper, journal article |
| | | | Project case study |
| | | | Policy recommendation |
| | | | Thematic working papers |
| | | | Rapid response activities e.g. coaching, mentoring, bulletins |
| | | | Seminars |
| | | | Conference |
| | | | 'live' presentations to audience or 'presentation pack' |
| | | | staff development activity |
| | | | staff development material |
| | | | facilitated workshop |
| | | | exhibition or visual display |
| | | | web site or web site pages |
| | | | other – please specify |

**Evaluation final reports**

The evaluator and the project manager – or whoever is commissioning the evaluation report should work through the following checklist as near to the beginning of the evaluation process as possible. Different decisions will have different cost and time implications and the evaluator needs to include these in the evaluation plan.

A) Format

Yes ☐ No ☐ Are there restrictions on length?

Is yes approximately how many pages or words?

Yes ☐ No ☐ Are there restrictions on style?

Please specify any restrictions on page size, house style, design and layout, stylesheets etc.

Yes ☐ No ☐ Will the report be commercially printed?

Yes ☐ No ☐ Are there print restrictions?

Please specify software programmes, presentation formats etc.

Who will produce the final version for the printers?

B) Content

Who are the target groups? (please specify)

Introduction
- Summary of background `problem' i.e. what 'gap in the market' was the project trying to fill.
- Aims and brief summary of project.
- Structure of the report.

Context
- Socio-economic / demographic profile of location (important if audience is not local).
- Institutional profile (important for audiences drawn from other sectors and locations).
- Infrastructure profile (may be important for transnational audiences).
- Background to 'subject area' i.e. key research, key issues (important for audiences who are from other disciplines).

Project Narrative
The story of the project from:
- Needs identification
- Aims and objectives
- Methodology and methods to final output.

Results
- Outcomes.
- Deliverables (what was achieved intentionally and why).
- Consequences or impacts (unintentional outcomes).

These could be tabulated as numeric details where is exists, followed by a discussion of the results.

Issues arising
- Issues arising from the project and lessons learned (against the dimensions of performance to be evaluated).

It is important that problem areas are identified and, as far as possible, what solutions were found. It is equally important to record effective practice and 'good news' and even more important to say why this happened.

Conclusions
- Summary of main findings
- Future work
  - What are the items still unresolved?
  - What are the 'gaps' still left unfilled?
  - What new areas of work / research are needed.
- Recommendations for practitioners, policy makers, promoters.

Appendices
Evaluation brief
Evaluation methodology and design
Timelines for project and evaluation
Details of data collection tools
Groups and individuals interviewed
Details of data analysis tools

**Michael Scrivens's Key Evaluation Checklist (KEC)**

A Standard Format For Program Evaluation Reports

PART A: PRELIMINARIES

A1. EXECUTIVE SUMMARY (Usually a short version of Checkpoints 11-15).
A2. PREFACE (Source & nature of request or need leading to the evaluation; identification of client for evaluation i.e., person officially commissioning it; adjustments to request after discussion (especially, is the conclusion to provide grading/ranking/apportioning; of merit/worth/significance); previous evaluations, if any; acknowledgments, etc.)
A3. METHODOLOGY (Investigative & data-analytic procedures used in this evaluation and their justification - the 'logic of the evaluation' (possibly requires lit review).

PART B: FOUNDATIONS

1. BACKGROUND & CONTEXT (Historical, contemporary, and projected settings; upstream stakeholders e.g., backers, staff, designers, funders; official rationale/program logic; related legislation/policy; lit review of similar programs).
2. DESCRIPTIONS & DEFINITIONS (Official description & correct and complete description of program, in enough detail to recognise, and if appropriate, replicate it; may include components, goals, targets, and the staff version of 'logic of the program;' note any patterns/ analogies/ metaphors; give meaning of technical terms if not known to likely audiences).
3. CONSUMERS (Comprises (i) impactees, i.e., the immediate recipients of services - sometimes called the 'clients' of the program, not of the evaluation; and (ii) those indirectly affected, whether targeted or not, through ‚ripple effects')
4. RESOURCES & ASSETS (For (i) the program (i.e., the actual plus potential ‚inputs'); for (ii) the consumer, and for (iii) other stakeholders). Includes quality of e.g., staff, not just quantity.
5. VALUES (Needs during as well as at start of program run; criteria of merit for a program of this kind; legislative requirements; definitional and descriptive accuracy standards; personal/ organizational goals; other supportable standards (for example, professional, legal, logical, ethical, scientific, market, historical and preferential when others do not supervene))

PART C: SUB-EVALUATIONS

6. PROCESS (Assessment of everything significant that happens before direct impacts on consumers occur: includes the evaluation of: goals, design, implementation (that is, degree of implementation of intended program), management, activities, procedures, results from early stages of process that do not directly impact consumers; merit of the official and staff logic of the program if not running goal-free).
7. OUTCOMES (Impact on consumers: direct/indirect, intended/unintended, immediate/short-term/long-term); short-term effects are sometimes

called outputs long term results are sometimes called effects, longer-term is durability.
8. COSTS (Money & non-money, direct & indirect, opportunity costs as well as actual costs, start-up/maintenance/upgrade/closedown; by time period, by payer; and by components if relevant and possible,…)
9. COMPARISONS (Identify other possible ways of getting about the same benefits from similar resources, the best of these are the 'critical competitors'; should usually rank and grade them all).
10. GENERALISABILITY (To other sites, types of program, times, staff, etc., that is, evaluate the potential export value).

## PART D: CONCLUSIONS

11. OVERALL SIGNIFICANCE (The sum of the above; focused on the consumers' needs, but including reference to the client's needs and wants; conclusion may be unidimensional (grading or ranking) or multi-dimensional (profiling).
12. EXPLANATIONS (Of successes/failures; in terms of optimal logic i.e., the best possible account; in terms of components and management; and only if it's appropriate to go beyond 11 and you have a defensible program theory.)
13. RECOMMENDATIONS Micro ones 'fall out' of the evaluation; macro ones require far more information than the evaluation, e.g., system constraints on funding decisions, as well as subject-matter, for example, a better program theory.
14. REPORTS (Audience reading; designing, justifying, and creating a way to convey the above conclusions: single or multiple, verbal or written, public or private, technical or non-technical or both, monomedia or multimedia).
15. META-EVALUATION (Evaluation of this evaluation: strengths/limitations/other uses: apply this list to the evaluation itself or use a special meta-evaluation checklist; or a meta-evaluator).

### Written Evaluation Reports

(From Reporting Evaluation Results of Extension Programs) C.L. Taylor and Larry R. Arrington. This is another tool we have found exceptionally useful.
In writing an evaluation report the following components are essential:
– Purpose of the evaluation.
– Method (procedure).
– Results.
Optional content would include:
– Background information.
– Expanded methodology.
– Conclusions.
The following are major points to be considered in writing the evaluative report:
1. Cover
– Use quality paper.
– Include: title, date, source.

2. Title Page
– Basically contains the same information as the cover.

3. Table of Contents
– Designed to help the reader understand the report.
– Include enough information to make it useful to the reader.
– Add sub-headings as appropriate.

4. Purpose
You might want to include:
– Goals or objectives, that is, what was the [project] expected to accom–lish? This information could come from the long range plans. In addition, what was the evaluation to assess?
– Accomplishment of objectives?
– Side effects?
– Wide range of outcomes?
– Background of the program(s) being evaluated:
– What is it?
– What could or should be?
– Significance of difference between what is" and „what could be".
– Problems identified.
– Profile of clientele.

5. Method
This might include:
– Population studied.
– The survey instrument.
– Data collection procedures.
– Data analysis.
Usually this is a brief section. It could also include evaluation design.

6. Results
– Display and discuss what you found.
– You may want to include „typical" quotes from respondents.
– Usually you would put most tables in the appendix where results are about different levels (input, activities, involvement, reactions, knowledge, practice change and end results) of [a project], arrange the discussion in the same hierarchy.
– Avoid evaluator bias in reporting results.

7. Conclusions
– Based upon the evidence that you collected, what conclusions can you draw from the data?
This section contains the agents' or evaluator's judgments about the results of the [project].
Conclusions:
– Focus on objectives of the evaluation.
– Deal with unanticipated results if they are significant.
– Evidence must be in the report to support the conclusions.
– Remember, it's important to let colleagues read and critique conclusions.
– Colleagues involved in the evaluation should agree on the conclusions.
If a written report is to be mailed to someone, it should have a cover letter explaining the evaluation.

Generally, any technical material should be in the footnotes. Technical material is that which may be complex or of little value to most of the consumers of the evaluation report. A formula to calculate sample size would generally be considered technical.

Use visual illustrations, charts, tables, and graphs where possible. However, they should not be expected to stand alone. Tables require written or oral explanation. The proverb "a picture is worth a thousand words," suggests that charts, graphs, and pictures add to evaluation reports by presenting the same material in a different format.

A summary or abstract may be placed at the end or beginning of the report. It's a brief overview of the report. It might contain the objectives of the [project], major findings of the evaluation, conclusions reached by the evaluator and in some cases, recommendations. The summary or abstract should contain enough essential information to enable a reader to obtain a picture of the program and its benefits or accomplishments in a brief period of time.

### The Oral Report

– When you report evaluation results orally, 'do what comes naturally' (Fitz-Gibbon, 1978).
– Be at ease. This implies that you know your audience.
– Practice your presentation and make it interesting by doing a variety of things other than merely talking. For example, use slides, transparencies, and role playing.
– Get your audience involved by letting them ask questions and predict results.
– Oral reporting may be either face-to-face or presented on radio and TV. Regardless of where the oral report is presented, be certain you are prepared.

The following suggestions are presented regarding the oral report.

– Do not read the report to the audience.
– Stay within time limits.
– Make 'team' presentations with presenters rotating back and forth.
– Communicate clearly.
– Briefly orient the audience to what you did and why you did it. Give emphasis to what you found, your conclusions, and your recommendations.
– Be enthusiastic and well prepared. Look like you enjoy what you are doing.
– Use visuals to help you communicate.
– Use percentages. They are meaningful to most people.
– Use good graphics.

### Mass Media Reporting

Rocheleau (1986) holds the opinion that the public is generally unaware of [project] evaluation findings. The main explanation for this lack of awareness is because evaluators and project professionals have not used the mass media to disseminate evaluation findings of educational [projects] (Taylor et al., 1992). However, using mass media to publicize the results of a [project] must be done with care. It is well documented that (Garber, 1980) the media like to focus on the negative or failure, especially concerning programmes of the government and other public agencies. Rocheleau's (1986) work suggests

that, when using the mass media to communicate the results of [projects] the following guidelines may be helpful:

– USE CONCRETE AND INDIVIDUAL CASES
  The media opposes the use of analytical, quantitative, or abstract aspects of an issue. It is generally accepted that news must be personalized to be interesting to readers. Goldberg (1975) says there is a push in the press to present situations affecting many people by focusing on one specific example. This allows reporters to deal in personalities rather than statistics. Success stories are a good way to communicate evaluation results of a [project].
– COMPASSION RATHER THAN COST-EFFECTIVENESS
  Graber (1980) reveals that formal [project] evaluation material and other types of analytical or statistical information concerning [projects] takes a back seat in most media coverage of program. In the case of educational programmes, evaluators and [other] professionals probably would have few occasions to use this approach. However, where appropriate, the sympathetic is more likely to be printed and read than the statistical or analytical information.

# Literature

**Recommended readings:**

Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). Program evaluation.: Alternative approaches and practical guidelines (2nd ed.). New York: Longman.

Stufflebeam, D. L. (1985). Systematic evaluation. Boston: Kluwer- Nijhoff.

Scriven,M. (1996). Types of evaluation and types of evaluator. Evaluation practice, 17, 151-162.

**References:**

Alkin, M.C. (1969) . Evaluation Theory development. Evaluation comment 2, 2-7.

Alkin, M.C. (1991) Evaluation Theory Development: II. In M.W. McLaughlin & D.C Phillips (Eds), Evaluation and Education: At quarter century. Year Ninetieth yearbook of the National Society for the Study of Education , Part II, Chicago: University of Chicago press.)

Alkin, M.C., & Solmon, L.C. (Eds). (1983). The cost of Evaluation. Beverly Hills, CA: Sage.

Alkin, M.C. Stecher, B.M. & Geiger, F.L. (1982). Title I Evaluation: Utility and Factors influencing use. Northridge, CA: Educational Evaluation Associates.

Shadish, D.L. Newman, M.A. Scheirer, & C. Wye (Eds), Guiding Principles for Evaluators. New Directions for Program Evaluation, no. 34, 19-26.

Baker, E.L., & Niemi, D. (1996). School and program evaluation. In D.C. Berliner & R.C. Calfee (Eds), Handbook of Educational psychology. New York: Macmillan.

Bell, J.B. (1994). Managing evaluation projects step by step. In J.S. Wholey, H.P.Hatry, & K.E. Newcomer (Eds), Handbook of Practical Program Evaluation. San Francisco: Jossey-Bass.

Bernstein, D.J. (1999). Comments on Perrin's "Effective use and misuse of performance Measurement." American Journal of Evaluation, 20, 85-94.

Bickman, L. (Ed) ,(1987). Using program theory in evaluation. New Directions for program Evaluation, No.33. San Francisco: Jossey Baass.

Bickman, L. (Ed.) (1990). Advances in program theory. New Directions in Program Evaluation, No. 47. San Francisco: Jossey Bass.

Bloom, B.S., Englehart, M. D., Furst , E.J., Hill, W.H., & Krathwohll, D.R. (1956) Taxonomy of educational objectives: Handbook I. Cognitive domain. New York: David McKay.

Brandon, P.R. (1998). Stakeholder participation for the purpose of helping ensure evaluation validity: Bridgeing the gap between collaborative and non- collaborative evaluations. American Journal of Evaluations, 19, 325-337.

Nowakowski, J.R. (1983). Program evaluation: A practitioners guide for trainers and educators. Boston: Kluwer-Nijhoff.

Byrk, A.S. (Ed). (1983). Stakeholder-based evaluation. New Directions for Program Evaluation, No. 17, San Francisco: Jossey Bass, 1983.

Canadian Evaluation Society. (1992). Standards for Program evaluation in canada: A discussion paper. Canadian Journal of Program Evaluation, 7, 157-170.

Chelimsky, E. (1997). The comming transformations in evaluation.In E . Chelimsky & W.R. Shadish (eds) Evaluation for the 21st century: A handbook. Thousand Oaks, CA: Sage,

Chelimsky, E., & Shadish, W.R. (1997). Evaluation for the 21st century: A handbook. Thousand Oaks, CA: Sage

Cook,D.L. (1996). Program evaluation and review technique: Applications in education (Monograph no.17). Washington., DC: U.S. Office of Education Cooperative Research.

Cook, T.D.(1997). .R.Shadish (Eds.). Evaluation for the 21st century: A Handbook. Thousand Oaks, CA: Sage

Cousins,J.B., & Earl, L.M. (1992). The case for participatory evaluation. Educational Evaluation and Policy Analysis, 14(4), 397-418.

Cronbach, L.JL (1982). Designing evaluations of educational and social programs. San Francisco: Jossey Bass.

Cronbach, L.J., Ambron, S.R., Dornbusch, S.M., Hess, R.D., Hornik, R.C., Phillips, D.C., Walker, D.F., & Weiner, S.S. (1980). Towards reform of program evaluation. San Francisco: Jossey Bass.

Datta, L. (1999). The ehtics of evaluation neutrality and advocacy. In J.L. Fitzpatrick & M . Morris (Eds.), Current and emerging ethical challenges in evaluations. New Directions for evaluation, No. 82, 77-88.San Francisco: Jossey Bass.

ETS Test Collection Catalogue. (1991). Phoenix, AZ: Oryx. Evaluation Research Society Standards Committee. (1982). Evaluation Research Society Standards for program evaluation. In P.H.Rossi (Ed.), Standards for Evaluation practice. New Directions for Program Evaluation, No.15, 7-19 . San Francisco: Jossey Bass.

Fetterman, D.M. (1984). Ethnography in educational evaluation. Beverly Hills, CA: Sage.

Fetterman, D.M. (1996). Empowerment evaluation: An Introduction to theory and practice. In D.M. Fetterman, S. Kaftairian, & A. Wandersman(Ed.), Empowerment evaluation: Knowledge and tools for self-assesment and accountability. Thousand Oaks, CA: Sage.

Fetterman, D.M.,Kaftarian, S., and Wandersman, A. (Eds.). (1996). Empowerment evaluation: Knowledge and tools for self-assesment and accountability. Thousand Oaks, CA: Sage.in evaluation. Preparation programs. In J.W. Altschuld & M. Engle (Eds), the preperation of professional evaluators: Issues, perspectives, and progarams. New directions for

Program evaluation, No.62, 41-50. San Francisco :Jossey Bass.

Fitzpartick, J.L. (1999). Ehtics in diciplines and professions related to evaluation. In J.L. Fitzpatrick & Morrsi (Eds), Current and emerging ethical challenges in evaluation. New Directions for program evaluation, No. 82. San Francisco, : Jossey Bass.

Green, J.C. (1987). Stakeholder Participation in evaluation design: Is it worth the effort? Evaluation and Program planning, 10, 379-394.

Greene, J.C., &Garcelli. V.J., (Ed) (1997) . Advances in mixed-method evaluation: The challenges and benefits of intergrating diverse paradigms. New Directions for program F., No. 74. San Francosco: Jossey

Guba, E.G., & Lincoln, Y.S. (1981). Effective Evaluation. San Francisco: Jossey Bass.

Guba, E.G., & Lincoln, Y.S. (1989). Fourth Generation Evaluation. Thousand Oaks, CA: Sage.

Guba, E.G., &Lincoln, Y.S. (1994). Competing paradigms.in qualitative research. In N.K. Denzine & Y. S. Lincoln (Eds.), Handbook of qualitative research. Thousand Oaks, CA: Sage.

Hackman, R., Wageman, R. (1995). Total Quality Management: Empirical, conceptual, and practical issues. Administrative Science Quarterly, 40, 309-342.

House, E.R.(1980). Evaluating with validity. Beverly Hills, CA: Sage.

House,E.R. (1983a). Assumptions underlying evaluation modles. In G.F. Madaus, M. Scrivean, & D.L.Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation. Boston: Kluwer-Nijhoff.

House, E.R. (1994b). Intergrating the quantitative and qualitative. In C.S. Reichardt & S.F. Rallis (Eds.), the qualitative-quantitative debate: New perspectives. New Directions for program evaluation, No. 61, 13-22. San Francisco: Jossey Bass.

House, E.R., & Howe, K.R. (1999). Values in evaluation and social research. Thousand Oaks, CA: Sage

Kirkpatrick, D.L.(1977). Evaluating training programs: Evidence Vs. proof. Training and Development journal, 31, 9-12.

Kirkpatrick, D.L. (1983). Four Steps to measuring training effectiveness. Personnel Administrator, 28, 19-25.

Lincoln, Y.S., & Guba, E.G. (1985). Naturalistic inquiery. Beverly Hills, CA: Sage

MacDonald, L.B. (1976). Evaluation of evaluation. Urban review, 7, 3-14.

MacDonald, L. B. (1976). Evaluation and the control of education. In D. Tawney (Ed.), Curriculum evaluation today: Trends and implications. Schools Council Research Studies Series. London: Macmillan.

Madaus, G.F., Scriven, M., & Stufflebeam, D. L. (Eds.), (1983). Educational modles: Viewpoints on educational and human services evaluation. Boston: Kluwer-Nijhoff.

Mark, M. M., & Shotland, R. L. (1985b). Stakeholder-based evaluation and value judgements: The role of perceived power and legitimacy in the selection stakeholder groups. Evaluation Review, 9, 605-626.

Mertens, D. M. (1994). Training evaluators: Unique skills and knowledge. In J.W. Altschuld and M. Engle (Eds.), The preparation of professional evaluators: Issues, Perspectives, and, programs. New Directions for Program Evaluation, No. 62 , 17-27. San Francisco: Jossey Bass.

Newman, D. L., & Brown, R. D. (1996). Applied ethics for program evaluation. Beverly Hills, CA: Sage.

Nillson, N., & Hogben, D. (1983). Metaevaluation. In E. R. House (Eds.), Philosophy of evaluation, No. 19, 83-97. San Francisco: Jossey Bass.

Patton, M. Q. (1975). Alternative evaluation research paradigm. Grand Forks: North Dakota Study Group on Evaluation.

Patton, M. Q. (1987). Utilization-Focused evaluation. (2nd ed .). Beverly Hills,CA: Sage.

Patton, M. Q. (1994). Developmental evaluation. Evaluation Practice, 15(3), 311-319.

Patton, M. Q. (1996). Utilization-Focused evaluation. (3rd ed.). Thousand Oaks, CA:
Sage.
Patton, M. Q. (1997b). Toward distinguishing empowerment evaluation and placing it in a larger context. Evaluation Practice, 18(2), 147-163.
Patton, M. Q. (2002). Qualitative evaluation and research methods (3rd ed.) Thousand Oaks, CA: Sage.
Provus, M. M. (1971). Discrepancy evaluation. Berkley CA: McCutchan.
Ridings, J. M. & Stufflbeam, D. L. (1981). Evaluation reflections: The project to Develop standards for educational evaluation: Its and future. Studies in Educational Evaluation, 7, 3-16.
Rossi, P. H., Freeman, H. E., (1998). Evaluation. A systematic approach (6th ed.). Newbury Park, CA: Sage.
Sanders, J. R. (1995). Standards and principles. In W. R. Shandish, D. L. Newman, M. A. Scheirer, & C. Wyre (Eds.) Guiding Principles for evaluators. New Directions for Program Evaluation, No. 66, 47-52. San Fransisco: Jossey Bass.
Sanders, J. R., & Cunningham, D. J. (1974). Techniques and procedures for formative evaluation. In G. D. Borich (Ed.), Evaluating educational programs and products. Englewood Cliffs, N. J: Educational Technology.
Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (Ed), curriculum evaluation. (American Educational Research Association Monogragh Series on Evaluation, No. 1, pp. 39-83). Chicago: Rand McNally.
Scriven, M. (1974a). Evaluation perspectives and procedures. In W. J. Popham (Ed.), Evaluation in education. Berkley, CA:McCutchan.
Scriven,M. (1996). Types of evaluation and types of evaluator. Evaluation practice, 17, 151-162.
Shadish, W. R. (1993). Critical multiplism: A research strategy and its attendant tactics. In L. Serchrest (Ed.), Program Evaluation: A pluralistic enterprise. New Directions for Program Evaluation, No. 60, 13-57. San Francisco: Jossey Bass.
Shadish, W. R. (1994). Need-based evaluation theory: What do you need to Know to do a good evaluation? Evaluation Practice, 15, 347-358.
Shotland, R. L., & Mark R. L. Shotland (1987). Improving inferences from multiple methods. In M. M. Mark & R. L. Shotland (Ed.) Multiple methods in program evaluation. New Directions for Program Evaluation, No. 35. San Francisco: Jossey Bass.
Stufflebeam, D. L. (1971). The relavence of the CIPP evaluation model for educational accountability. Journal of Research and Development. In Education, 5, 19-25.
Stufflebeam, D. L. (2001a) Evaluation checklists: Practical tools for guiding and D.
Stufflebeam, D. L. (2001b) Evaluation models. New Directions for Evaluation, No. 89. San Francisco: Jossey Bass
Stufflebeam, D. L. (1985). Systematic evaluation. Boston: Kluwer- Nijhoff.
United Ways Of America. (1996). Measuring program outcomes. Alexandria, VA:United Way of America.
Weiss, C. H. ( 1997). Theory-based evaluation: Past, present & future. In D. Rog and D. Fournier (Eds.), progress and future directions in evaluation:

Pespectives on theory, practice and methods. New Directions for Evaluation, No. 76. San Francisco: Jossey Bass.

Wholey, J. S. (1994). Assessing the feasibility and likely usefulness of evaluation. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), Hadbook of practical program evaluation. San Francisco: Jossey Bass.

Wholey, J. S. (1996). Formative and summative evaluation: Related issues in performance measurement. Evaluation practice, 17, 145-149.

Worthen, B. R. (1975). Competencies for educational research and evaluation. Educational Researcher, 4, 13-16.

Worthen, B. R. (1994). Is evaluation a mature profession that warrants the preperation of evaluation profesionals? In J. W . Altschuld, and M. Engle (Eds.), The Preperation of professional evaluators: Issues, Perspectives and programs. New Directions for program evaluation, No. 62, 3-15. San Francisco: Jossey Bass.

Worthen, .B. R., & Sanders, J. R. (1987). Educational Evaluation: Alternative approaches and practical guidelines. New York: Longman.

Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). Program evaluation.: Alternative approaches and practical guidelines (2nd ed.). New York: Longman.