

# The Use of Formative Evaluation by Government Agencies

Strategic Evaluation Working Paper

First Posted on 3 June 2004 as Version 1-2-4

When citing this document the following format can be used:

Duignan, P. (2004, June 3). *The Use of Formative Evaluation by Government Agencies* [WWW document]. From Strategic

Evaluation Web Site. URL

<http://www.strategicevaluation.info/se/documents/121pdf.htm>

1.

Dr Paul Duignan<sup>1</sup>

Parker Duignan Ltd

<sup>1</sup> Dr Paul Duignan works for Parker Duignan Ltd Wellington and is a Senior Research Fellow at the Centre for Social and Health Outcomes Massey University Auckland, New Zealand (SHORE).

If using models from this paper in any context please source them to this document,  
the reference format is provided on the front page.

## Contents

|     |   |    |
|-----|---|----|
| 1   | Executive Summary .....   | 4  |
| 2   | Evaluation Terminology.....                                     | 5  |
| 3   | Evaluation Purposes .....                                       | 7  |
| 3.1 | Formative Evaluation.....                                       | 8  |
| 3.2 | Process Evaluation.....   | 10 |
| 3.3 | Impact/Outcome Evaluation.....                                  | 11 |
| 4   | The use of formative evaluation.....                            | 14 |
| 4.1 | Evaluation Planning: Strategic use of evaluation resources..... | 14 |
| 4.2 | Setting up formative evaluation.....                            | 18 |
| 4.3 | Capacity for undertaking formative evaluation .....             | 22 |
| 4.4 | Increasing acceptance and use of formative evaluation .....     | 23 |
| 5   | References.....   | 24 |

## Executive Summary

This paper outlines the use of formative evaluation as a key aspect of evaluation planning within government agencies. The first section of the paper clarifies evaluation terminology in some detail and distinguishes between three *purposes* or *types* of evaluation - formative, process and impact/outcome evaluation. Formative evaluation is concerned with optimising programme planning and implementation; process evaluation with describing the context and course of a programme; and impact/outcome evaluation with assessing the outcomes of a programme. Evaluation planning needs to use a four step process to determine the right mix of evaluation types for any evaluation: 1) developing overall sector evaluation priorities; 2) considering outcome evaluation options; 3) considering formative evaluation options; 4) considering process evaluation options.

The use of formative evaluation can be set up in an organisation through: 1) capacity building in formative evaluation techniques; 2) formative evaluation peer review systems; 3) an internal formative evaluation team; 4) an external formative evaluation team; 5) use of specific formative evaluation models for hard to evaluate programmes. Capacity building for formative evaluation needs to take place in regard to: 1) decision makers awareness of the potential of formative evaluation; 2) institutional arrangements and values to support independent formative evaluation; 3) the development of formative evaluation specialists. Increasing the acceptance and use of formative evaluation could be achieved by: 1) gaining acceptance of the concept of formative evaluation by key decision makers; 2) developing appropriate formative evaluation skills in staff at all levels; 3) setting up one or more pilot projects to evaluate the use of formative evaluation for programme improvement; 4) drawing on the findings from the pilot to routinely include formative evaluation amongst potential range of evaluation types.

## 1 Evaluation Terminology

There is considerable diversity and confusion in the field of evaluation regarding evaluation terms, particularly amongst stakeholders considering questions related to evaluation, research and evaluation priorities. This confusion is exacerbated when evaluation terms from different conceptual levels are used in discussions about evaluation. For instance, there can be a discussion about how an evaluation should proceed where one stakeholder is talking about an *approach*, e.g. empowerment evaluation; another about a *purpose* of evaluation, e.g. outcome evaluation; a third about a *method*, e.g. a survey; and a fourth about a *design*, e.g. a quasi-experimental design. In order to clarify the use of evaluation terms four conceptual levels of evaluation terminology can be identified (Duignan 2003) as follows:

- Evaluation *approaches*: the overall philosophical and value orientation to the task of undertaking evaluations (e.g. goal-free, kaupapa Maori, utilisation focused, empowerment, strategic evaluation).
- Evaluation *purposes* (sometimes called *types*): the purpose of an evaluation (or part of an evaluation) (e.g. evaluability assessment, formative, process, impact/outcome, summative evaluation).
- Evaluation *methods*: the specific research or other methods (for data collection and analysis) used in evaluations (e.g. surveys, interviews, focus groups).
- Evaluation *designs*: the way in which the *methods* of an evaluation are used to answer evaluation questions for a particular *purpose* under the overall framework of an evaluation *approach* (e.g. quasi-experimental, case study).

The following table lists some examples of common terminology used under the four conceptual headings: approaches, purposes, methods and designs:

| <b>Approaches</b>   | <b>Purposes</b>                                | <b>Methods</b>   |
|---|--|--|
| Utilisation-focused evaluation                            | Formative Evaluation:                          | Stakeholder consultation   |
| Empowerment evaluation                                    | - <i>Evaluability assessment</i>               | Literature & other programme review/prospective evaluation synthesis   |
| Stakeholder evaluation                                    | - <i>Design</i>                                | Objective setting and programme (implementation) logic critique  |
| Goal-free evaluation                                      | - <i>Developmental</i>                         | Evaluation workshops/hui   |
| Naturalistic, constructivist or 4th generation evaluation | - <i>Formative</i>                             | Pre-testing and piloting   |
| Theory Based Evaluation                                   | - <i>Implementation</i>                        | Data collection (written, photos, video or audio recording). These can be archival, administrative /routine records collection/records collected specifically for the evaluation |
| Kaupapa Maori research/evaluation                         | Process Evaluation                             | Observation (passive/environmental audit/participant)  |
| Positivist (quasi) experimental evaluation                | - <i>programme description</i>                 | Document analysis  |
| Strategic evaluation                                      | - <i>programme documentation</i>               | Interviews, questionnaires, feedback sheets (key informant/ participant)   |
|   | - <i>programme origin and context analysis</i> | Focus groups   |
|   | Impact/Outcome Evaluation:                     | Each of these methods may involve the following methods of analysis  |
|   | - <i>Impact</i>                                | - <i>Expert assessment</i>   |
|   | - <i>Outcome</i>                               | - <i>Quantitative / statistical analysis (e.g. statistical testing, regression, global scaling)</i>  |
|   | - <i>Summative</i>                             | - <i>Qualitative analysis (e.g. thematic analysis, discourse analysis)</i>   |
|   |  | - <i>Economic analysis (e.g. cost benefit, cost effectiveness analysis).</i>   |

## Designs

An evaluation design is how the approaches, purposes and methods are combined in a particular evaluation. This may include using one of the “designs” identified in the literature such as: case study designs (e.g. *illustrative, exploratory, critical instance, programme implementation, programme effects, cumulative, success case method*) or experimental and quasi-experimental designs (e.g. *non-intervention control group with pre-test and post-test, removed intervention design with pre-test and post-test, repeated intervention design, regression-discontinuity design*). In addition, an overall evaluation design using either case study or quasi-experimental designs could well include a mix of formative, process and outcome evaluation elements (*purposes*) and a range of *methods* within an approach. In addition, complementary case study and quasi-experimental designs can be used in the same evaluation.

This paper focuses on one aspect of the *purpose* or *type* of evaluation – formative evaluation. The next section explores in more detail the concept of evaluation *purpose*.

## 2 Evaluation Purposes

There are various ways of describing the *purpose* of any evaluation activity, e.g. evaluability assessment, design, developmental, formative, implementation, process, impact, outcome and summative evaluation. The evaluation *purpose* is best understood as identifying the *use* to which evaluation activity is going to be put. Recent years have seen moves to develop types of evaluation that are of use right across a programme lifecycle.

The range of evaluation purposes are described in various ways in the literature. One way of thinking about them is to cluster them into three broad evaluation *types* or *purposes* as follows:

- Formative evaluation – evaluative activity designed to improve the design, development, formation and implementation of a programme. (including prospective evaluation synthesis, evaluability assessment, design, developmental, formative, implementation evaluation)
- Process – evaluation to describe what happened in the course and context of the programme. It is conceptually useful to limit the scope of the term process evaluation otherwise - because the term process could conceivably cover all of a programme from its inception to its outcomes - it could encompass the entirety of evaluation.
- Impact, outcome and summative – to assess the impact and outcome of a programme, and (in the case of summative, make an overall evaluative judgment about the worth of a programme).

It should be noted that any particular evaluation method can have more than one purpose. The three broad purposes of evaluation are discussed in more detail below.

## 2.1 Formative Evaluation

Formative Evaluation includes the terms – *evaluability assessment / design / developmental / formative / implementation* evaluation – it is any evaluation activity directed at improving a programme's design, planning, development and implementation. Formative evaluation is a disciplined approach to ensuring that a programme is well developed. It is a somewhat later development in the history of the evaluation discipline (McClintock 1986). There are various models for how it can be undertaken, but it is directed at ensuring independent constructively critical input into programme development. For instance, a separately funded independent formative evaluation team can work alongside programme planners. This team

critically assesses the decisions that are being made in the programme and can provide regular, formal feedback to programme planners and programme funders (Dehar, Duignan et al. 1993). Formative evaluation may use both quantitative and qualitative methods. Another activity which can be seen as a sub-set of formative evaluation is *evaluability assessment* (Wholey 1987) which ranges from assessment of a programme simply to make the decision as to whether or not an impact/outcome evaluation should be undertaken, through to more complex formative evaluation-type processes such as looking at, and assisting with developing and modifying the programme (intervention) logic of the programme. An additional term related to formative evaluation is *prospective evaluation synthesis* (U.S. General Accounting Office 1990). This method, which has been proposed by the General Accounting Office in the United States, starts before a programme is launched and looks at the results from previous evaluation studies, develops a possible programme (intervention) logic and assesses the likelihood of such a programme being successful on the basis of the previous evaluation research and the nature of the programme logic.

The preparation, critique and use of programme (or intervention) logic is a major element in formative evaluation. Programme logic sets out the detailed steps of the way in which it is believed a programme will achieve its outcomes. Once a programme logic has been developed it can be used for a number of important purposes:

- As the basis for peer review of the programme by outside experts who can critique the links which are assumed between activities and programme outcomes
- The improvement of the programme through laying bare the reasoning behind it
- Considering alternatives strategies which could be used in the programme to achieve the same results

- The risks to achieving programme outcomes which need to be managed
- Identifying intermediate outcome measures (outcomes hierarchy) which, when measured, show the programme achieving progress towards its final outcomes.

## 2.2 Process Evaluation

Process Evaluation is any evaluative activity directed at describing or documenting what actually happened in the course or context of a programme. Its first use is in communicating best practice to others who want to replicate elements of a successful programme. For instance, taking as an example a *Programme X* that has been shown to be effective from an impact/outcome evaluation. Just knowing that *Programme X* is effective is, in itself, insufficient for someone elsewhere wanting to replicate the programme. Process evaluation gives outside stakeholders detailed information on what was done, what problems arose and what solutions were adopted in the programme.

A second use of process evaluation is in the interpretation of impact/outcome evaluation results. For instance, a programme may not have proved successful on outcome evaluation. However, when looking at the process evaluation from the programme, it will be clear that the negative outcome was a result of specific events that derailed the programme (for example the coordinator leaving half-way through programme implementation). In the light of this, one would not dismiss the possibility that this type of programme, if implemented as intended, could be effective.

A third use of process evaluation is when it is used to examine the context of a programme and the decision making leading up to that programme being introduced. For instance there may be "problem definition creep" in the early decision making phase about a programme's objectives and the type of programme that should be run. This may lead to a programme being designed which is "easy to implement" rather than one which is more difficult to implement but which would

be more likely to achieve success with the substantive problem being addressed (Duignan and Casswell 1989).

### 2.3 Impact/Outcome Evaluation

Impact/Outcome evaluation also includes the terms *intermediate outcome* evaluation and *impact* evaluation. Impact evaluation is used in evaluation in at least two ways, firstly simply being another word for outcome evaluation, secondly it is sometimes used to describe measuring shorter-term impacts of a programme in contrast to a programme's longer term outcomes. Impact/outcome evaluation is any evaluation activity directed at determining the positive or negative, intended or unintended, intermediate or longer-term outcomes of a programme. It is sometimes also referred to as summative evaluation, which also includes the aspect of making an overall assessment of a programme as a whole. Outcome evaluation looks at whether a programme has achieved the outcomes it is seeking. Where such evaluation can be done, it provides very useful information for stakeholders, particularly if it is in a form in which the effectiveness of the programme can be compared with that of other programmes directed at achieving the same outcomes.

In real world programmes, the final outcomes being sought by a programme may take a number of years to achieve and may be outside the timeframe of the measurements being undertaken in an evaluation. Given this, there is a way in which evaluation designers can give outcome-type results earlier within time frames that are more useful for policy decision-making. This entails the development of an *outcome hierarchy* for a programme or policy based on the programme (intervention) logic. This is a set of outcomes that range from immediate outcomes of the programme or policy through to intermediate and then final outcomes. An argument needs to be mounted that each step in the outcomes hierarchy is likely to imply that the next step will occur. If this argument is sound then intermediate outcomes can be measured at an earlier level, within reasonable time frames, and the assumption made that there is a good chance that the later steps in the outcome hierarchy will also take place in due course.

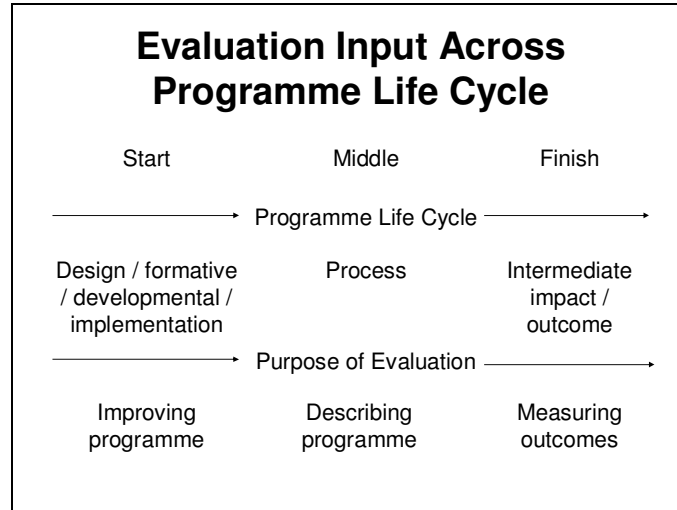
In the past, outcome evaluation has tended to be largely quantitative and this generally remains the case. However, this notion is now being challenged and the argument mounted in that qualitative outcome evaluation is feasible for a range of programmes (Mohr, 1999).

There can often be tension between different stakeholders as to the purpose they are seeking from an evaluation. For instance, a central policy maker is naturally very interested in an impact/outcome purpose for an evaluation because they want clear and 'objective' comparative information on which to make resource allocation decisions. For example they are seeking clear conclusions from an evaluation such as, *Programme X has been shown on outcome evaluation to be three times more effective than Programme Y while costing the same amount per participant.* On the other hand, programme management and staff are often more interested in formative evaluation for programme improvement and external stakeholders may also be interested in process evaluation so that they know exactly what happened in a programme in addition to their interest in impact/outcome evaluation.

The three purposes for evaluation can be approximately related to three stages in the programme life-cycle, the start, middle and finish of a programme.

Diagram 1 illustrates the relationship of the three different purposes of evaluation to programme stages.

**Diagram 1: Evaluation "Purposes" or "Types"**



These three purposes of evaluation are obviously linked – information arising out of a formative evaluation can be used for both improving a programme and also as part of process evaluation documenting what happened in a programme. Formative, process and outcome evaluation from one programme can be used to feed into the formative evaluation of a new programme.

The three-way typology of evaluation purposes proposed here is preferred by the author for its practical usefulness in a number of situations which arise in normal evaluation practice. There are other typologies of evaluation purposes which are obviously neither right nor wrong from a theoretical point of view, just more or less useful in practice. It is useful to briefly contrast the three-way division (formative, process and intermediate impact/outcome) with several alternative typologies of evaluation purposes.

The first of these alternatives is where a two-way distinction is made between formative and summative evaluation. Just using this two-way distinction loses the emphasis in the three-way split on process evaluation as describing or documenting what happened in the course and context of a programme. Having a clear and separate term for process evaluation can be useful in some evaluation situations. A second two-way distinction is sometimes used by evaluators who just distinguish

between process and outcome evaluation. In this typology process evaluation includes the evaluation activity covered under the heading of formative evaluation. This two-way distinction has the downside that it does not clearly identify formative evaluation as a legitimate form of evaluation. Raising the profile of formative evaluation is particularly important at the current time when a number of funders and other stakeholders still see evaluation as being just about impact/outcome evaluation. This is leading to under-optimised programmes being subject to expensive outcome evaluations when the priority is for formative evaluation.

### **3 The use of formative evaluation**

#### **3.1 Evaluation Planning: Strategic use of evaluation resources**

Using the three-way distinction between evaluation purposes – formative, process and outcome – enables decisions to be made about the allocation of evaluation resources right across a programme's lifecycle. This is crucial in situations where there are limited resources available for evaluation, such as in New Zealand. In an ideal situation a set of steps will be taken in deciding to allocate evaluation resources for the evaluation of any programme. These steps are set out below under the headings of evaluation strategy, outcome, formative and process evaluation:

##### **Step 1: Developing overall evaluation priorities**

*This step consists of prior definition of sector priority evaluation questions through stakeholder discussions (as outlined in Duignan (2004)) and an assessment of whether the evaluation of this project can answer any of the priority evaluation questions (if not then only low level administrative monitoring of the project should take place).*

Notes on this step: If there is no higher-level evaluation strategy, in a situation of limited resources, it is impossible to sensibly allocate evaluation resources to a particular programme evaluation. What happens in practice is that evaluation managers and evaluators working on an individual programme make decisions

about overall evaluation priority setting which may or may not be based on full information about sector priorities. Alternatively, a simplistic assumption is made at the highest level that an outcome evaluation will be methodologically possible, affordable and feasible without any exploration of whether or not such an outcome evaluation can actually be successfully achieved for the programme in question.

### **Step 2: Considering outcome evaluation**

*This step consists of an assessment of whether an outcome evaluation is:*

- *methodologically possible (propose design)*
- *able to be undertaken within the available budget*
- *feasible given the level of organisational and political support for the proposed design*
- *likely to actually get through to completion and yield useful results*

Notes on this step: If the evaluation of the programme outcomes has been identified as a priority in Step 1 following the assessment in step 2, a decision needs to be made regarding whether an outcome evaluation is going to be attempted. This needs to be then further reviewed in the light of the answers to step 4 below.

### **Step 3: Considering formative evaluation**

*This step consists of an assessment of whether programme:*

- *planning is likely to be sufficient to optimise the programme for success (if not then there is usually little point in undertaking a high-cost outcome evaluation)*
- *implementation is likely to be sufficient so as to optimise the programme for success (if not then there is little point in undertaking a high-cost outcome evaluation)*

Notes on this step: In most cases of large programmes some formative evaluation work needs to be done, even in order to answer the questions in this step. The decision to undertake an outcome evaluation (Step 2) needs to be reviewed in the light of the answer to the questions set out in this step which will be answered by a formative evaluation process.

#### **Step 4: Considering process evaluation**

*This step consists of an assessment of what process evaluation needs to be done in order to document the course and context of the programme in order to:*

- *interpret any outcome evaluation findings*
- *better understand the process of the programme*
- *enable sharing of best practice*
- *analyse and document the wider context of the development of the programme in order to ensure that more general lessons arising from the attempt to mount this particular programme are learnt.*

Notes on this step: Process evaluation can be used to meet any of the needs listed in this step. Even in the situation where an outcome evaluation has been ruled out, it is too late for formative evaluation, and programme implementation is thought to be poor; a process evaluation of the context of the programme can point to the problems which have resulted in this situation and assist in ensuring that they are not endlessly repeated in the future.

In contrast to the careful strategic consideration of evaluation options as outlined in the steps above, there is often an unquestioned assumption from the highest levels down that a programme should just be subject to an outcome evaluation. This arises from a natural assumption when taking a high-level policy view of evaluation that outcome evaluations should be undertaken on all priority programmes in order to determine whether or not such programmes work. The belief is that once this has been done the results of such outcome evaluations can be used to make decisions about which programmes should or should not continue to receive funding.

Unfortunately such an assumption only works in the following situations:

- it is methodologically possible and not too expensive to undertake outcome evaluations

- outcome evaluations can attribute the change in outcomes to the activity of a specific programme
- there is the political and institutional will (at all levels) to support evaluation designs which can provide rigorous information on causality (e.g. only providing services to an intervention group)
- the outcomes being measured are singular, or a small group of outcomes which are relatively easy to define to the satisfaction of all key stakeholders
- the outcomes will be achieved within relatively short timeframes (to enable their measurement in outcome evaluations)
- the entry and exit costs for providers providing the services being evaluated is low.

In reality the situation is often very different from this:

- because of the point in the project life-cycle (e.g. early in roll-out) when evaluators are brought in, or for other reasons, it is not methodologically possible and it is too expensive to undertake outcome evaluations
- because of a multitude of programmes with overlapping objectives in a “joined-up” social policy world it is not possible to attribute the change in outcomes to the activity of a specific programme
- there is not the political and institutional will (at all levels) to support evaluation designs which can provide rigorous information on causality (e.g. withholding services to an intervention group)
- the outcomes being measured are multiple and there is little consensus amongst stakeholders regarding their definition
- the outcomes will be achieved over long timeframes outside the timeframe of any feasible outcome evaluation
- the entry and exit costs for providers providing the services being evaluated is high with expensive capacity building having had to take place.

In the second situation, it is often more strategic to spend scarce evaluation resources on formative evaluation rather than attempt outcome evaluations which are doomed

to failure. Formative evaluation in these situations can maximise the chances of an intervention being successful. There are usually unrealistic expectations by key stakeholders about the feasibility of outcome evaluations and inexperienced (or pressurised) evaluation managers and evaluators end up undertaking unrealistic outcome evaluations which return little or nothing on the investment which has been made in them. In these situations, a careful strategic evaluation analysis would point to robust formative evaluation being the evaluation priority rather than outcome evaluation. This is not to say that formative evaluation should always be used in preference to outcome evaluation – it is just up until now formative evaluation has not tended to be considered in evaluation planning in those situations where it would be more appropriate. Undertaking the analysis in the steps above will ensure that the option of formative evaluation is carefully considered in evaluation planning.

Those involved in making the assessments required by the steps set out in this section need to have a high level of skills in the assessment of evaluation methodologies and be familiar with the realities of undertaking evaluations in the sector in which they are working.

Following the above four steps should lead to establishing an evaluation plan for a programme which has the right mix of formative, process and outcome evaluation; alternatively it may lead to a decision to not undertake any significant evaluation of a particular programme and simply use routine administrative monitoring of the programme.

### **3.2 Setting up formative evaluation**

There are various possibilities for setting up formative evaluation processes within an organisation. Some of the activity which takes place under the name of formative evaluation could take place under other headings such as: comprehensive programme planning, quality assurance, peer review and evidence-informed practice. It does not matter what name is used to describe such activity as long as it

takes place. It is up to the organisation concerned to decide how it will categorise such activity; however systems need to be in place to ensure that the key aspects of formative evaluation are undertaken regardless of what they are called.

One useful way of conceptualising formative evaluation is to see it as analogous to the process of gaining a warrant of fitness for a motor vehicle. A mechanic inspects the motor vehicle, using a check-list to see if the vehicle meets all of the specified requirements. If the vehicle does meet the requirements, then the mechanic does not need to be involved any more; however, in many cases a vehicle will need further work and the mechanic will offer to undertake such work. For large and complex programmes there is likely to always be work which needs to be undertaken by formative evaluators.

A key aspect to formative evaluation is that institutional arrangements need to be put in place to ensure that the checking of programme implementation which is being undertaken by the formative evaluator is independent of the programme which is being developed. If this is not done then the power of formative evaluation is lost.

An organisation can foster formative evaluation in a range of ways. These are set out below:

1. Capacity building in formative evaluation techniques

As part of comprehensive capacity building for research and evaluation (not dealt with in this paper) appropriate training can be provided to programme management and staff regarding formative evaluation techniques they can use to improve programme planning. For instance, objective setting, intervention logic development, and peer-review of evaluation plans. They can then use these in regard to programmes they are running or managing themselves. Alternatively, this capacity building may be undertaken under the heading of project planning. It has a formative evaluation element if institutional arrangements can be put in place which

encourage staff to take a critically independent view of the work that they are doing in programme planning.

## 2. Formative evaluation peer review systems

Once capacity has been built for staff to understand and be able to undertake appropriate formative evaluation tasks, peer-review formative evaluation systems can be set up so that staff can review each other's programme planning from a formative evaluation perspective. These systems need to be institutionalised so that they can provide protect the independence of the peer-review which is taking place.

## 3. Internal formative evaluation unit

For a larger organisation an internal formative evaluation unit can be established, the role of the unit would be to undertake formative evaluation of projects. It also requires institutionalised protection to ensure its independence.

## 4. External formative evaluation teams

For larger projects, under the spotlight from various stakeholders and being introduced under tight time frames, the only reliable way of ensuring independent formative evaluation is use an external formative evaluation team. This is where a formative evaluation team is contracted to work alongside the programme as a *critical friend* to the programme. This type of process has not been widely used in New Zealand. It is a different role from the central agency monitoring and auditing role which already takes place in regard to programmes being implemented by government agencies.

An example of this approach was used in the case of the Heartbeat New Zealand programme. This programme, funded out of the Health Vote was run by the NZ Heart Foundation. A group of formative evaluators, the Alcohol and Public Health Research Unit at the University of Auckland was commissioned to work alongside

the programme. The formative evaluation team assisted the programme in programme planning, intervention logic design, strategy selection and process and outcome evaluation design. The formative evaluation team reported on a regular basis to the funder on the progress the programme was making. Such an approach deals with one of the major problems in public sector programme planning – the demand to set up programmes in unrealistic time periods. This usually means that insufficient planning has been carried out. This affects the adequacy of programme implementation. The presence of a formative evaluation team working alongside the programme can ensure that it is being well planned and implemented and, at the same time, assure stakeholders that any additional planning time being requested by the programme planners is actually needed in order to ensure a well-formed programme.

#### 5. Use of formative evaluation models for hard to evaluate *centrally funded-locally implemented* programmes

A further use of formative evaluation is in the case of *centrally funded-locally implemented* programmes. In these cases those locally implementing the programme have considerable autonomy as to how it is implemented. Formative evaluation models have been developed which are designed to address the particular challenges associated with such programmes. The first, and most basic, issue for an evaluator looking at such programmes is *which programme objectives should the programme be evaluated against?* – those of the central funder or those of the individual localities which are implementing the programme. There are also multiple other challenges from such programmes such as how to inject information about evidence-based practice into the discussions at the locality level about the strategies which will be employed; how to get standardised reporting on programme implementation; and in some cases, how to get adequate reporting at all. A formative evaluation process, the Collaborative Ongoing Formative Evaluation Workshop Process (COFE), has been developed in order to address these evaluation challenges. Duignan (2002) outlines this process.

### 3.3 Capacity for undertaking formative evaluation

There are three areas in which formative evaluation capacity needs to be built.

#### 1. Decision makers awareness of the potential of formative evaluation

Many decision makers are unaware of the potential of formative evaluation to ensure that programmes are optimised for success. This is the first area of capacity which needs to be addressed. If decision makers do not know of the value that can be added by robust formative evaluation processes they will not employ it in regard to the programmes and activities under their control.

#### 2. Institutional arrangements and values to support independent formative evaluation

For formative evaluation to work effectively there need to be institutional arrangements to allow formative evaluators the independence to report frankly on the status of programme development. Such reporting can take a number of forms both written and verbal and take place in a number of forums. Under all government administrations, programme implementation is often an area which is contested by stakeholders; as a consequence there need to be robust institutional arrangements to ensure the independence of formative evaluation processes. Just as organisational culture has been developed over many years to accept the need for rigorous financial audit procedures, there needs to be the development of a public sector culture supportive of the different role of formative evaluators as a *critical friend* to a programme in its developmental and implementation stages.

#### 3. Formative evaluators skills

Evaluation is a relatively underdeveloped discipline in New Zealand and as a consequence New Zealand evaluators come from very diverse backgrounds and with different levels of skill. Formative evaluation is a relatively new concept in evaluation although some New Zealand evaluators have

pioneered the use of formative evaluation techniques. In addition, formative evaluation of large scale public sector programmes which are under the spotlight demands a range of highly developed evaluation, communication, strategic and diplomatic skills. As a consequence of this, finding evaluators who can undertake formative evaluation as envisaged in this paper is not necessarily easy.

### **3.4 Increasing acceptance and use of formative evaluation**

In the light of the discussion above, the development of formative evaluation needs to be a staged process. These stages are as follows:

1. Gain acceptance of the concept of formative evaluation by key decision makers within an organisation
2. Plan to develop appropriate formative evaluation skills in staff at all levels in the organisation either as part of evaluation capacity building or under the heading of comprehensive project planning
3. If there have not been some successful examples of formative evaluation within the organisation in regard to important types of programmes, set up one or more pilots using formative evaluation with large scale project (these pilots would include careful attention to the institutional arrangements by which the formative evaluators would report; careful selection of the formative evaluators; and a small-scale evaluation of the way in which the formative evaluation worked to improve the programme)
4. Drawing on the lessons learnt from the formative evaluation pilots to include formative evaluation as one of the types of evaluation which can be selected when planning for a programme evaluation .

## 4 Conclusion

This paper has discussed the use of formative evaluation by government agencies. Any evaluation planning needs to consider using the option of using some formative evaluation as part of a comprehensive evaluation strategy. An institutionalised and adequately resourced approach to formative evaluation has the potential to optimise the chances of success for many government programmes.

## 5 References

- Dehar, M., P. Duignan, & Casswell, S. (1993). Formative and process evaluation of health promotion and disease prevention programmes. *Evaluation Review* 17: 204-220.
- Duignan, P. and S. Casswell (2002, October 10) [Updated 14 May 2004] *Collaborative Ongoing Formative Evaluation Workshop Process (COFE) for Implementing Change* [WWW document]. From Strategic Evaluation Web Site. URL <http://www.strategievaluation.info/se/documents/107pdff.html>.
- Duignan, P. (2003). Approaches and terminology in programme and policy evaluation. *Evaluating Policy and Practice: A New Zealand Reader*. N. Lunt, C. Davidson and K. McKegg. Auckland, Pearson Education.
- Duignan, P. (2004, June 3). *Linking Research and Evaluation Plans to an Organisation's Statement of Intent (SOI)* [WWW document]. From The Strategic Evaluation Web Site. URL <http://www.strategievaluation.info/se/documents/120pdff.html>.
- Duignan, P. and S. Casswell (1989). Evaluating community development programs for health promotion: Problems illustrated by a New Zealand example. *Community Health Studies* XIII(1): 74-81.
- McClintock, C. (1986). Towards a theory of formative program evaluation. *Evaluation Studies Review Annual* 11: 205-223.
- Mohr, L. B. (1999). The Qualitative Method of Impact Analysis. *American Journal of Evaluation* 20(1): 69-84.

U.S. General Accounting Office (1990). *Prospective evaluation methods - the prospective evaluation synthesis* [WWW document]. URL

[http://www.gao.gov/policy/10\\_1\\_10.htm](http://www.gao.gov/policy/10_1_10.htm), Program Evaluation and Methodology Division.

Wholey, J. (1987). Evaluability assessment: developing program theory. *New Directions for Program Evaluation* 33: 77-92.

© Paul Duignan 2004