RURAL INDUSTRIES RESEARCH
& DEVELOPMENT CORPORATION

Review of

Evaluation in Agricultural Extension

by Jessica Dart, R John Petheram, Warren Straw
Institute of Land and Food Resources, Vic

Rural Industries Research and Development Corporation
Human Capital, Communications & Information Systems
Research and Development

RIRDC Publication No. 98/136 RIRDC Project No. VCA-3A

© 1998 Rural Industries Research and Development Corporation.
All rights reserved.

ISBN 0 642 57803 6
ISSN 1440-6845

"Review of Evaluation in Agricultural Extension"
Publication No 98/136
Project No VCA-3A

The information contained in this publication is intended for general use to assist public knowledge and discussion and to help improve the development of sustainable industries. The information should not be relied upon for the purpose of a particular matter. Specialist and/or appropriate legal advice should be obtained before any action or decision is taken on the basis of any material in this document. The Commonwealth of Australia, Rural Industries Research and Development Corporation, the authors or contributors do not assume liability of any kind whatsoever resulting from any person's use or reliance upon the content of this document.

This publication is copyright. However, RIRDC encourages wide dissemination of its research, providing the Corporation is clearly acknowledged. For any other enquiries concerning reproduction, contact the Publications Manager on phone 02 6272 3186.

Researchers Contact Details
Jessica Dart and R.J. Petheram
Institute of Land and Food Resources
Department of Crop Protection
RMB 3000
HORSHAM VIC 3401

Phone: 03 5362 222
Fax: 03 5362 2213

RIRDC Contact Details
Rural Industries Research and Development Corporation
Level 1, AMA House
42 Macquarie Street
BARTON ACT 2600
PO Box 4776
KINGSTON ACT 2604

Phone**:** 02 6272 4819
Fax**:** 02 6272 5877
Email: rirdc@rirdc.gov.au
Website: http://www.rirdc.gov.au

First published in November 1998
Printed on environmentally friendly paper by Union Offset

Foreword

Greater emphasis on accountability in the delivery of extension programs has put new focus on techniques for evaluating such programs.

This report discusses the discipline of program evaluation and then reviews five main forms of evaluation and illustrates them with case studies drawn from agricultural extension in Australia.

It is hoped that improved access to sound information on evaluation methods will be reflected in better planning of agricultural extension projects, more effective monitoring and evaluation of programs, and hence more efficient use of industry funds and greater benefits for clients of all research and development programs in rural Australia.

This report belongs to RIRDC's range of over 250 research publications and forms part of our Human Capital, Communications and Information Systems R&D program, which aims to enhance human capital and facilitate innovation in rural industries and communities.

Peter Core
Managing Director
Rural Industries Research and Development Corporation

Acknowledgments

Firstly we would like to thank all the people who contributed to the Review by sending us evaluation reports and answering our inquiries. All contributors are listed in Appendix 4. Special acknowledgment must be given to the practitioners who conducted the evaluations that make up the case studies in this report. Without their work we would have no concrete examples of evaluation to illustrate the various approaches to evaluation that have been taken in agricultural extension today.

Secondly we would like to thank and acknowledge the support of Dr. Barrie Bardsley for his timely advice and input into the Review. His balanced critique of early drafts greatly influenced the final product.

We would also like to thank John Owen, whose book *Program Evaluation: Forms and Approaches* greatly helped us to understand the various approaches to evaluation from the discipline of program evaluation *per se*. Also a big thank you to Jerome Winston from RMIT for his advice and encouragement to a group of non-professional evaluators exploring the vast field of evaluation.

Contents

FOREWORD III

ACKNOWLEDGMENTS I

EXECUTIVE SUMMARY IX

1. AIM, EMPHASIS AND REVIEW 1

1.1. Scope of the Review 1

1.2. Using this Text 2

1.3. Taking a Broad View of Evaluation 2

1.4. Time-Out for Some Definitions 3

2. EVALUATION IN AGRICULTURAL EXTENSION 6

2.1. Historical Changes in the Position of Evaluation in Extension 6

2.2. Relevance of Evaluation to Agricultural Extension Today 7

2.3. New Requirements and Challenges for Evaluation 9

2.4. Why is Formative Evaluation Important in Extension Programs? 11

3. PROGRAM EVALUATION 13

3.1. Relevance of Program Evaluation as a Discipline to Extension 13

3.2. Definition of Program Evaluation 13

3.3. Development of Program Evaluation as a Discipline 14

3.4. The Development of Evaluation in Australia 17

4. FORMS OF PROGRAM EVALUATION 18

4.1. Owen's Five Forms of Evaluation 18

4.2. Alternative Meta-Models for Classifying Evaluation 19

4.3. The Domains of Evaluation 19

4.4. Different Levels of Program Evaluation 21

5. CHARACTERISTICS OF EVALUATION IN AGRICULTURAL EXTENSION 22

5.1. Aim and Method 22

5.2. Criteria Used to Characterise the Range of Evaluation Studies 23

5.3. Some Results and Discussion 24

6. CASE STUDIES OF DIFFERENT FORMS OF EVALUATION IN AGRICULTURAL EXTENSION 29

7. FORM 1: EVALUATION FOR IMPACT ASSESSMENT 30

7.1. Goal-Based Impact Evaluation 30

7.2. Needs-Based Impact Evaluation 32

7.3. Financial Analysis of Outcome 35

7.4. Impact Evaluation for Illumination 39

8. FORM 2: MONITORING AND EVALUATION FOR PROGRAM MANAGEMENT 43

8.1. Performance Indicators Used in M&E 45

8.2. Evolutionary Approach to Project Monitoring 49

8.3. Quality Assurance 56

9. FORM 3: PROCESS EVALUATION 59

9.1. Process Evaluation Carried Out to Understand Program Delivery 59

9.2. Process Evaluation that is Particularly Responsive to Needs of Program Clients (farmers) 64

10. FORM 4: EVALUATION FOR DESIGN CLARIFICATION 68

10.1. Program Logic 68

11. FORM 5: EVALUATION FOR PROGRAM DEVELOPMENT 70

11.1. Economic Evaluation for Program Synthesis 70

11.2. Needs Assessment 70

11.3. Rapid Rural Appraisal 70

12. FRAMEWORKS FOR PLANNING PROGRAM EVALUATION 74

12.1. Bennett's Hierarchy 75

12.2. The Logical Framework 77

12.3. The Snyder Approach to Evaluation 81

12.4. The Three Rings of Extension Evaluation 84

12.5. Integrated Frameworks That Combine Several Forms of Evaluation 85

13. SUMMARY GUIDE TO APPROACHES AND METHODS IN EVALUATION 88

14. REFERENCES 94

15. GLOSSARY 100

MAIN ACRONYMS USED 102

APPENDIX 103

Appendix 1. Program Participation Index 103

Appendix 2. Evaluation Participation Index 103

Appendix 3. List of Case Studies and their Contributors 104

Appendix 4. List of Contributors 105

List of Figures, Tables and Case Studies

Figure 1 Diagrammatic representation of a purchaser–provider model 8
Figure 2 Some trends in focus of agricultural extension and implications for evaluation 11
Figure 3 Domains of evaluation for a hypothetical program 20
Figure 4 Distribution of the evaluation studies according to Owen's five forms..... 25
Figure 5 Distribution of evaluation studies according to the level of investigation according to Bennett's hierarchy 26
Figure 6 Level of participation in evaluation by farmer clients in the studies 27
Figure 7 Graph of net present value at different adoption rates 38
Figure 8 Information collection flow 41
Figure 9 Flow chart of activities in establishment of a QA program for Farm Pride Food..... 58
Figure 10 Evaluation strategy for a hypothetical program that employs multiple forms..... 74
Figure 11 The three rings of extension evaluation 84

Table 1 Owen's five forms of program evaluation..... 18
Table 2 Summary of main characteristics of the sample of evaluation studies 24
Table 3 An example of the summary table preceding each case study..... 29
Table 4 Example of matrix for recording M&E data for a rice development project over three years (shading indicates monitoring periods) 43
Table 5 Bennett's hierarchy 44
Table 6 Global assessment scale describing community groups 63
Table 7 Bennett's hierarchy with examples..... 76
Table 8 The logframe matrix: questions to be answered when filling in each cell of a logframe. 78
Table 9 Example of a simplified synergy matrix for the cotton industry 85
Table 10 Strategic frameworks to conduct evaluations for a program 89
Table 11 Conceptual models of evaluation..... 90
Table 12 Methods of evaluation..... 91

Case Study 1 Landcare: does it make a difference? 31
Case Study 2 A case study undertaken with the West Hume Landcare Group as part of an international research project on Participatory Watershed Development..... 33
Case Study 3 Cost–benefit analysis of the Target 10 dairy extension program 36
Case Study 4 Evaluation of a participatory process project in Western India..... 40
Case Study 5 IPM program for potato pests: an example of monitoring & evaluation..... 47
Case Study 6 FARMSCAPE for improved management of dryland cropping..... 51
Case Study 7 An evolutionary approach to facilitating organisational learning: An experiment by the Christian Commission for Development in Bangladesh..... 53
Case Study 8 Quality assurance for Farm Pride Food, Victoria 57
Case Study 9 National Whole Catchment Management: review and analysis of processes..... 60
Case Study 10 Corangamite Salinity Program and global assessment scaling..... 62
Case Study 11 Improving the usefulness of the Pigpulse Report, an evaluation..... 65
Case Study 12 A workshop to refine the program logic for Coranagmite Salinity Program..... 68
Case Study 13 Understanding farmer decision-making on land use 72
Case Study 14 A logframe for proposed project to support Landcare group activities in the Wimmera River Catchment 79
Case Study 15 Evaluation of a program to replace imprisonment using the Snyder model..... 81
Case Study 16 Target 10 framework for project evaluation 86

Executive Summary

This publication presents the findings of a review of program evaluation in the field of agricultural extension in Australia. Over 100 documents were reviewed from a wide range of sources, ranging from rural extension centres, to universities and agriculture departments in various states, private consultants and individuals.

Without doubt there are some very able and experienced evaluators active in the field of program evaluation in agriculture, and the case studies in this document reflect the diversity and creativity present in this field. However, during the course of the Review many extension agents and managers expressed a lack of knowledge, skills and confidence in this area. Also, much of the existing evaluation practice in agriculture is based around a few rather limited approaches and methods such as mail survey and cost–benefit analysis, with little reference to developments and trends in program evaluation in other disciplines. Extension managers are generally aware of this deficiency and are seeking appropriate training for extension staff, to enable them to meet the increasing need for evaluation for various purposes in extension programs. It was in the context of this perceived need for training materials that the Review was conceived and sponsored by RIRDC in 1996/97.

A revealing feature of the work for the Review was the difficulty experienced in finding well-documented examples of certain forms of program evaluation, such as formal systems of monitoring and evaluation (M&E) in agricultural extension in Australia. However, the published literature on evaluation covers just one part of the 'lived experience' of applied evaluation: only the information that authors choose to write about comes into the public domain; other potentially valuable information often remains unrecorded and hidden. It could well be that practitioners do not fully realise the value of their experiences for others in similar situations.

The Review is structured around the 'form' (or purpose) of evaluation. Five main forms of evaluation from Owen 1993) are described and case studies are included as examples of each from agricultural extension. The Review concerns evaluation at the level of 'programs' (or projects), and not at the level of small events (seminars, meetings, courses) or at 'higher' levels of extension policy evaluation. Other texts are available to guide the extension agent in evaluating small events, and policy level evaluation is seen as outside the normal duties of the field extensionist. Much of the background information and theory of evaluation presented here, however, should be useful to practitioners at any level of evaluation. Methods of evaluation are not a focus of the Review, but many methods are mentioned in the case studies and in summary tables in Section 13.

The literature on program evaluation *per se* is vast, full of jargon and often confusing. An attempt is made in Section 1 to define some important terms and to explain some of the main 'domains' of evaluations covered in this literature, ie models, frameworks, and methods of evaluation. The Glossary provides definitions of other common terms used in evaluation. In Section 2 an effort is made to summarise some of the history and growth of program evaluation as a discipline throughout the world, as this has relevance to its status in Australia and in extension today.

After reviewing the material available, the authors' impressions are that the most common approach to evaluation in agricultural extension in Australia is summative evaluation, using both qualitative and quantitative data. Evaluations were usually conducted while the program was in a settled stage of development and were carried out by external evaluators. The clients of the program (usually farmers) were not generally involved in developing the indicators for the evaluation. The most common aim of evaluation was to provide a report to justify spending and to understand whether the stated objectives of the program had been met. There were strong indications that the range of approaches to evaluation in agricultural extension is expanding. A number of new and innovative evaluations are under way that are more formative and qualitative than has been the norm in the past.

Different readers may wish to progress through the Review via different routes: the table of contents identifies particular topics. Those new to evaluation should first seek to understand the different purposes for which evaluation is used, and to familiarise themselves with the main terminology of the discipline. Once the reader has decided which purpose (or form) of evaluation is particularly pertinent to their situation, they may wish to read the section dealing with that particular form, and the case studies presented to illustrate the form. Ideas on methods may arise from the case studies, and may be followed up by referring to the tables of methods in Section 13 and then in the evaluation literature.

The 15 case studies in the Review are intended to provide examples of each main form of evaluation, but also to provoke thought about ideas for developing appropriate evaluation approaches for particular contexts. In the course of conducting the Review many approaches to evaluation were discovered from other disciplines that are potentially valuable for agricultural extension in Australia. The reader new to evaluation is encouraged to work with this and other texts and with experienced evaluators to assess the purpose of evaluation and the most appropriate forms and methods of evaluation for their needs. The literature quoted in the Review should lead all levels of extension practitioner to new and interesting ideas to incorporate into evaluation practice.

1. AIM, EMPHASIS AND REVIEW

The aim of this publication is to review the current forms of evaluation used in agricultural extension in Australia. The Review entailed collecting information on over 100 evaluations from the field of agricultural extension and categorising these according to the 'purpose' of evaluation and other criteria. The main body of the Review comprises descriptions of various forms of evaluation found in agricultural extension today. The forms are described with regard to the purpose of the evaluation, and are supported by case studies taken from agricultural extension to illustrate each form. In addition to reviewing evaluation by purpose, case studies that describe frameworks for developing built-in evaluation strategies for projects (or programs) are also presented.

The Review is also intended to introduce the reader to the broader concept of 'program evaluation', its origins and its potential applications for agricultural extension. Therefore, the study has made use of literature from disciplines outside agricultural extension: on program evaluation *per se*, on health and education programs, and on monitoring and evaluation of agricultural development projects overseas. One intention of the Review is to provide a source of materials that can be used for training agricultural extension agents and other rural professionals in the field of program evaluation.

1.1. SCOPE OF THE REVIEW

Scriven (1991) states that evaluation is about determining the merit or worth of the object being evaluated. The object can be a program, a project, a product, a policy, or a one-off event. The Review concentrates on the evaluation of programs or projects in the field of agricultural extension. However, even within one program, evaluation can be made up of several domains. Firstly a program may have an evaluation strategy; which could comprise several forms of evaluation, each serving different purposes and operating to different time frames. Each of the different forms of evaluation may employ various methods of evaluation. This document is mainly concerned with reviewing the various forms of evaluation, which are functions of the evaluative purpose. For a description of the various domains of evaluation the reader is referred to Section 4. Because many terms are used ambiguously in the vast literature on evaluation, a list of definitions is provided in Section 1.4 and numerous terms are covered in the Glossary.

One of the most important points that has come out of the Review is that an evaluation should start with a close examination of the purpose of the evaluation and a clear understanding of the target clientele. It is not until the purpose and key evaluation questions have been agreed upon that the selection of appropriate methods should be considered. Thus we have concentrated on the various forms of evaluation, which are related to the purpose of the evaluation.

1.2. USING THIS TEXT

Different readers may wish to select different starting points and routes in using this document. The main themes of each Section are listed below:

- Section 1 — Introduction to the range and scope of evaluation. Taking a broad view of evaluation. Some main concepts and terms in evaluation.
- Section 2 — The general state of play of evaluation in agricultural extension.
- Section 3 — An outline of the discipline of program evaluation, its origins and relevance for agricultural extension in Australia.
- Section 4 — Different approaches to evaluation and some meta-models for classifying and describing these approaches.
- Section 5 — Summary findings from an analysis of 50 evaluation reports with regards to approach, type and methods of evaluation.
- Sections 6 -11 — Five common forms of evaluation, with case studies of evaluations from agricultural extension in Australia:
  - Form 1 Evaluation for Impact Assessment
  - Form 2 Monitoring and Evaluation for Program Management
  - Form 3 Process Evaluation
  - Form 4 Evaluation for Design Clarification
  - Form 5 Evaluation for Program Development
- Section 12 — Frameworks used for planning integrated evaluation.
- Section 13 — Summary of methods and literature sources.

1.3. TAKING A BROAD VIEW OF EVALUATION

At first glance the literature on program evaluation is daunting: the texts are full of jargon and messages are often contradictory. Worthen *et al.* (1997) state that nearly 60 different proposals for the conduct of evaluation have been developed. Each model of evaluation has its own built-in assumptions. Evaluation literature is vast and fragmented and generally aimed at evaluation theorists rather than practitioners. There is a plethora of methods and tools of evaluation, and a confusing amount of literature and different opinions as to the best model of evaluation to follow.

Anecdotal evidence from discussions with people from various organisations during this study indicates that there is a lack of confidence among extension workers as to how to design evaluations, select appropriate tools and utilise findings. It appears that there is a tendency to stick to particular approaches and to a handful of familiar methods, and that this may result from a lack of understanding of evaluation in its broad sense.

Historically, much of the debate in program evaluation was concerned with the merits/disadvantages of one method against another (Chen 1990). For example, the debate between qualitative and quantitative methods and the debate between naturalistic and experimental methods. Chen argues that there are major problems associated with a methods-oriented approach. Firstly, an adherence to any one approach may cause an evaluator to lose sight of the multiple domains and issues inherent in an evaluation, and secondly traditional evaluation perspectives (such as the experimental paradigm) tend to develop evaluation principles and strategies with little reference to the context of the evaluation. He argues that any one method is only valid under particular circumstances.

A major thesis of the Review is that good evaluation is not merely about learning a few methods; it is more about understanding the concepts of program evaluation and being able to pick and choose from a wide range of tools to meet particular needs.

1.4. TIME-OUT FOR SOME DEFINITIONS

It is difficult to write a review of evaluation without using much of the jargon that makes the literature of program evaluation so difficult to digest. While every attempt is made to minimise the use of jargon, there are a few terms that need definition and agreement before going any further.

Evaluation and monitoring

There are many definitions of evaluation and some of these are discussed later. Put simply, evaluation involves **determining the worth or merit of whatever is being evaluated** (Scriven 1991). Many different uses can made of those value judgements, from assessing the financial or social impact of a program, to improving program designs or planning new programs. The term **monitoring** also appears to have many definitions and uses. In this text, the term relates to a management context in which monitoring is a process of systematically collecting information in a consistent manner to service basic information needs.

Formative and summative evaluation

Scriven (1967) was the first evaluator to write about the differences between formative and summative evaluation. Since then the terms have become almost universally accepted in the field of evaluation. The distinction between formative and summative evaluation is concerned with the ultimate use of the value judgement. **Formative evaluation** is conducted to provide program staff with judgements useful in *improving the program*. **Summative evaluation** is generally conducted after completion of the program (or when a program has stabilised) and for the benefit of some external audience or decision-maker. The findings from a summative evaluation could be used to decide whether to continue a program or not, or to justify program spending. The main difference is that the aim of a summative evaluation is to report *on* the program, whereas a formative evaluation reports *to* the program (Scriven 1991).

Process evaluation and outcome evaluation

Process evaluation focuses on the process of a program rather than the outcome and is generally carried out during the implementation of a program. It involves the collection of information about program activities and about the delivery of a program. Process evaluation is generally formative in nature, but can also be summative, for example when the stakeholders need to know about the actual delivery of a program for justification of program spending. **Outcome evaluation** is carried out towards the end of a program's life and is aimed at assessing the program to establish whether the desired outcomes have been reached. Outcome evaluation is generally summative, but may be formative when the findings are used to improve future or other existing programs.

Goal-based and needs-based evaluation

Goal (or objective) **based evaluations** are carried out to judge whether the stated goals or objectives of a program have been achieved. Tyler (1967) was among the first to develop and use goal-based evaluation. In this approach goals are taken as given and decisions about the success of a program are based on the extent to which the goals are reached. According to Owen (1993) the main tasks of planning and implementing a goal-based evaluation are to identify the real goals of the program and to decide on how to determine whether the program has led to the attainment of goals. A standard procedure in goal-based evaluation is to develop outcome measures that have strong face validity. However, it is sometimes necessary to use substitute measures that stand in the place of preferred or ideal measures.

Scriven (1967) discussed the limitations of goal-based evaluation and used the term 'goal-free' evaluation. His argument was that testing the extent to which the goals of a program have been achieved does not determine the worth of that program for society in general; it does not assess the goals themselves. Goal-free evaluation is a form of 'needs-based' evaluation. In needs-based evaluation the evaluator does not concentrate solely on the stated objectives of a program but also evaluates the impact of the program with regard to the needs of the stakeholder group and the needs of society in general.

Quantitative and qualitative data analysis

Quantitative data analysis in evaluation usually refers to approaches involving heavy use of numerical measurement and data analysis methods, from social science or from accounting (e.g. cost–benefit analysis). It focuses on the testing of specific hypotheses that are smaller parts of some larger theoretical perspective. **Quantitative research** emphasises standardisation, precision, objectivity and reliability of measurement. Quantitative research is characterised by generating data that can be analysed statistically (Scriven 1991).

Qualitative data analysis in evaluation refers to the part of evaluation that cannot be usefully reduced to quantitative measures. A large part of contemporary evaluation is qualitative, meaning that description and interpretation make up most of it (Scriven 1991). The use of **qualitative methods** can provide evaluators with rich, first-hand information on questions such as how a program is implemented, the patterns of interaction between stakeholders, the kind of day-to-day problems that are confronted

by program staff, and so on. Qualitative inquiry emphasises ‘thick description’, ie obtaining ‘real’ data that illuminate everyday patterns of actions. It also tends to focus on social processes rather than outcomes (Worthen *et al*.1997).

Readers are referred to the Glossary for further definitions of terms commonly used in evaluation literature.

2. EVALUATION IN AGRICULTURAL EXTENSION

Extension refers to the conscious use of communication of information to help people form sound opinions and make good decisions (Van den Ban and Hawkins 1996). The evaluation of agricultural extension programs implies the systematic collection of information about the activities, characteristics, and outcomes of a program to make judgments about the program, improve its effectiveness, and/or inform decisions about future programming.

2.1. HISTORICAL CHANGES IN THE POSITION OF EVALUATION IN EXTENSION

The practice of agricultural extension in Australia has been influenced from the very beginning by the environment in which farmers and extension agencies have had to operate. A key element in this environment has been government policy. In the early days of extension, the policy emphasis was on increased production, both to provide for domestic requirements and to produce increased income from exported produce. Later, the emphasis shifted to productivity (increasing output per unit of scarce input) rather than production *per se*. To achieve these policy objectives, extension moved through three identifiable phases: one, the provision of straight technical information; two, the provision of information taking into account financial effects of advice; and three, consideration of the whole farm business (Dillon 1965). Arguably, the last of these stages has never been completed. In addition, had Dillon been writing in the 1980s or 1990s he would undoubtedly have referred to a fourth stage, consideration of the environment. Dillon's perspective is that of the farmer, but what of the government's perspective, as the body providing much of the funding for extension?

While developments in government policy as it related to extension varied between the States, in the 1960s and early 1970s the emphasis was fairly clearly on providing an educational service for farmers to help them solve their problems. In such a setting, evaluation, where it was conducted at all, tended to emphasise providing feedback so that the process better met farmers' needs. Evaluation was much discussed at extension conferences, and national workshops were conducted to focus attention on taking a professional approach to evaluating the efficiency and effectiveness of extension. Issues such as how to measure behaviour change were addressed (Riches 1973). In the optimism of the period extension workers felt valued by their farmer clients, and most often felt that they were too busy helping these clients to spend time on formal evaluation procedures (Bardsley, B. 1997 pers. comm).

During the 1970s, government policy began to change. Labor governments at State and Federal level began to demand more accountability for the expenditure of public funds, and in most cases, to see extension as a tool to bring about those changes sought by government rather than as an educational process. This changed the perspective for evaluation, since governments began to ask what effect extension was having in terms of new technology being applied, and whether extension led to faster adoption. Governments also wanted to know what effects on production and productivity were being achieved by the extension of new and improved management and technology. Much of the evaluation conducted at this stage was based upon assessing the achievements of institutional goals or of individual advisers (Bardsley, 1997 pers. comm).

2.2. RELEVANCE OF EVALUATION TO AGRICULTURAL EXTENSION TODAY

Several factors combine to make evaluation an important issue in agricultural extension programs today. The changing face of agricultural extension itself, the squeeze on budgets, and the climate of accountability for program spending all contribute to the increasing role that evaluation has to play. The following paragraphs cover factors that have led to evaluation becoming a pertinent issue in the 1990s.

I. The move towards the user pays principle

Governments worldwide have traditionally taken the financial responsibility for providing agricultural extension services. The prime reasons have been to increase agricultural production or to improve the efficiency of agricultural production (Lees 1991). Both reasons are concerned with increasing output to provide more food, fuel and fibre to meet growing domestic or export demand. Watson *et. al* (1992), in a *Review of Field Based Service*, recommended that the Victorian department should concentrate on areas of market failure where the private sector is unwilling or unable to provide services to farmers. A Victorian state departmental corporate plan gave the overall rationale for the continuation of government extension services as: "to support increased income from primary production within sustainable agricultural systems" (Department of Agriculture 1993).

However, the global climate of budget deficits has led governments increasingly to question the public funding of agricultural extension services and the efficiency of such expenditure in terms of returns to society. At the Australian Agricultural Extension Conference in Brisbane in 1987, the question of funding government extension services received considerable attention. The majority of the discussion concerned the application of 'user pays' as this was seen as the most effective way to overcome the decline in extension budgets.

Lees (1991) pointed out that given the increasing tendency for governments to look for ways of decreasing public spending on agricultural extension services, there is a growing need in Australia to conduct evaluation which provides evidence as to whether:

- public sector involvement in agricultural extension is justified *per se*
- the overall level of expenditure on agricultural expenditure is justified
- the allocation of funds amongst alternative competing extension activities is optimal
- the expenditure on particular extension activities is efficient and effective
- the greatest economic and social benefit is obtained by the use of these funds on agricultural extension.

The relative importance of agriculture in the economic growth of industrialised countries has declined. Added to this, increased use of externally-purchased inputs has changed the nature of publicly funded extension services and led to a questioning of the rationale for the delivery of such services by governments (Cary 1993). Public funding for research and development (R&D) is shrinking and agricultural programs must meet increasingly high standards of accountability and quality. Well-designed evaluation methods are seen as an essential basis for planning and implementing high

quality agricultural extension programs, and hence assuring public benefits from R&D expenditure.

II. Outcome oriented programs and the move to purchaser–provider models

In all states of Australia there is a movement towards the purchaser–provider model for government-provided agricultural extension programs. In Victoria this new model was introduced in July 1997. Figure 1 illustrates how the Victorian purchaser–provider model is implemented within the Primary Industries purchasing group. This approach aims to increase impact, decrease overhead costs, to improve evaluation, amongst other things. The model also has a strong outcomes focus: these outcomes are broadly defined at the outset of a project, then later in the project cycle they are defined more specifically and to be specific, measurable, accountable, realistic and timebound (SMART). The projects are focused on the achievement of outcomes and must have a built-in evaluation strategy. To obtain tenders, providers will have to move towards projects with a built-in evaluation component (Agricultural Industries Team 1997).

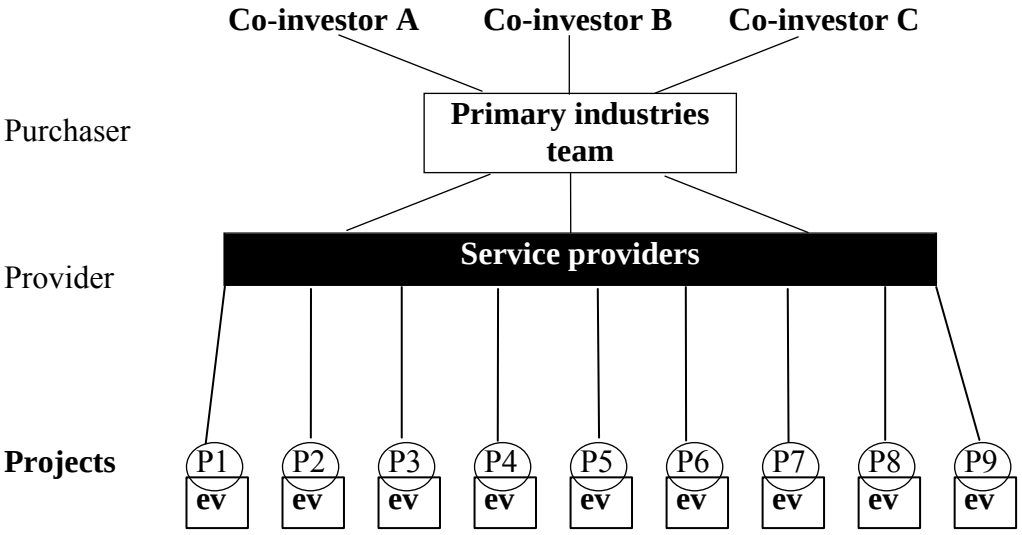


Figure 1 Diagrammatic representation of a purchaser–provider model

The bottom line in Figure 1 indicates that evaluation (ev) is to be built into every project. The purchasers are expected to define outcomes for industry. It may soon be the case that projects without clearly defined outcomes and built-in evaluation strategies will not obtain funding. This once again underlines the increasing importance of evaluation in agricultural extension. It also explains the pressure felt by extension agencies to undertake 'impact' (or outcomes) focused evaluation.

III. Increasing demand for training in evaluation by rural professionals

As a result of the changing nature of agricultural extension at the macro level, there is a strong demand from Government departments, R&D Corporations and other institutions for personnel who are well-trained in evaluation to take on a wide range of evaluation tasks in rural industries. In 1995, during a series of workshops aimed at gaining feedback on how to improve the performance of agricultural extension in

Victoria, extension workers listed evaluation skills and techniques more frequently than any other issue needed for their professional development (Straw, W. 1997 pers. comm).

The current lack of expertise in program evaluation in the field of agricultural extension is perplexing to those who are aware that historically many important contributions to program evaluation have been made from the field of agricultural extension (e.g. the theory of diffusion of innovation). Existing information on evaluation is largely out of date and inaccessible to many extension agents. Perhaps this can be explained by the nature of the aim of agricultural extension itself, and by the physical science-oriented training of most extension agents. Agricultural extension aims to change behaviour through the use of communication. Behaviour change is commonly perceived to be difficult to measure, especially in quantitative terms. As most agricultural extension workers have little training in social science they are likely to be unfamiliar with the range of qualitative data collection methods.

This is not to say that extension practitioners and program developers are not *aware* of the need to evaluate. Evaluation is high on the agenda of extension organisations (Straw *et al.* 1997). Commodity research groups, such as the Dairy Research and Development Corporation make it a condition that all new projects have a built-in evaluation component.

2.3. NEW REQUIREMENTS AND CHALLENGES FOR EVALUATION

Increasingly complex issues

The information/knowledge system (IKS) of agricultural extension in Australia is becoming increasingly complex. The shift in focus of extension from production orientation to a more holistic view, including: sustainability, improved marketing, increasing attention to the management of the farm business and maintaining profitability under declining terms of trade, adds to this complexity. There are also more actors operating in the capacity of 'knowledge bearers' to farmers than in the past. The implication of this increasing complexity of the IKS on evaluation is that it is becoming more difficult to distinguish the impact of one program or actor from that of another (Coutts 1997). Evaluation studies need to take into consideration the implication of the integrated impact of the workings of the IKS on the rural sector.

The new culture of agricultural extension

Worldwide, agricultural R&D programs are becoming more participatory in focus. New extension theories and methods are gaining acceptance, such as those of Chambers (1983) who argued that research activities should begin and end with the farmer. In Australia similar concepts emerged during the 60s and 70s (Williams 1968 and Bardsley 1981). Today the new emphasis is on adult learning, understanding existing farming systems, reflective practice of extension workers as facilitators and a stronger focus on evaluation. To cater for this change in the culture of extension practice, new evaluation methods will have to be developed or adapted from other disciplines. These methods need to be participatory, creative, formative, empowering and aid decision making. This type of evaluation is generally known as **formative**

evaluation. **Summative** (or outcome) evaluation is also important but is generally needed for more conventional purposes of accountability and justification of funding, for which there is also an increase in demand.

Increasingly participatory, group oriented R&D - and evaluation

There is now considerable agricultural R&D activity that was instigated on the basis of claims that farmer participation is critical to the generation of technologies that are relevant to farmers (Bebbington *et al.* 1994, Foale 1997). This is another aspect of the new nature of agricultural extension, which has led to requirements for different approaches to evaluation. Allen (in press) states that because participatory programs are designed to be responsive to changing community needs, one of the most pressing challenges is to develop participatory and systems-based evaluative processes to allow for ongoing learning, correction and adjustment by all parties. Alternative methods are needed for evaluating new participatory, group-oriented approaches to extension; in other words, participatory programs require participatory evaluation. Some of these changes in the focus of extension and their effect on evaluation are depicted in Figure 2.

The adoption of philosophies of farming systems research at most international agricultural research institutes has had a large impact on national programs in many countries (Anderson and Dillon 1985). In Australia systems based models of R&D have been recommended or applied in some regions (see McCown 1991; Clark *et al.* 1996; Foale 1997) although the approaches vary widely. The Hawkesbury model of agricultural systems development (Bawden 1991; Macadam 1996) is based on experiential learning (Kolb 1984), which encapsulates the concept of learning through experience. Most such systems approaches incorporate a strong element of iterative evaluation.

Overall, program evaluation is now more relevant to agricultural extension in its many forms than it has ever been. As explained, this trend has been gaining momentum over recent years. One would therefore expect extension professionals to have developed interest and expertise in program evaluation in response to these clear signals. Apparently this has not generally been the case.

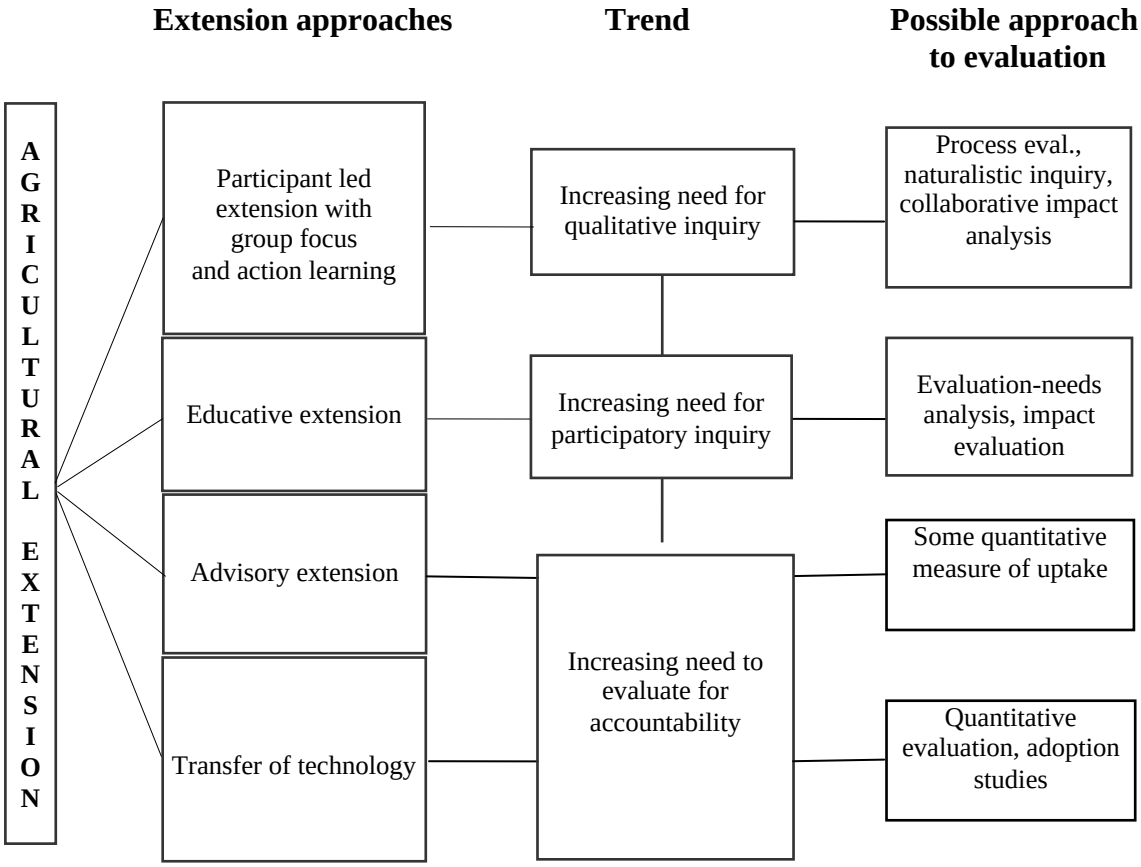


Figure 2 Some trends in focus of agricultural extension and implications for evaluation

2.4. WHY IS FORMATIVE EVALUATION IMPORTANT IN EXTENSION PROGRAMS?

While a focus on outcome evaluation is understandable in the context of budget restrictions and an increasing need to be accountable for public spending, there are several arguments that highlight the danger of focusing exclusively on outcome (summative) evaluation. The need for more formative evaluation in agricultural extension underpins the larger movement towards increasingly 'process'-oriented programs in general. The term 'process-orientation' in this context is used to describe programs that do not have rigidly defined goals at the onset of the program but have a defined purpose and a set of potential goals. Programs with a 'process' orientation are difficult to evaluate as they are carried out in variable, unpredictable situations; they produce outputs that are hard to measure objectively and often have permeable boundaries and less-than-direct relationships between inputs and outputs (Farrington and Nelson 1997).

The modern literature on program evaluation stresses various '**dangers**' of sticking to rigid outcome evaluations, including:

- Outcome evaluation carried out at the end of a program may fail to provide information on how the outcomes were achieved
- In the case of attitude change (frequently the goal of agricultural extension programs) the causal factors for change may be difficult to identify
- There is danger in programs with a strong focus on outcome evaluation that the indicators become the mission
- Outcome oriented evaluation built into the program model may drive the model into concentrating on producing measurable outcomes when perhaps the most desirable outcome is not easily measurable!

Some points mentioned on the **strengths** of formative (or process) evaluation are that:

- Knowledge gained about the process that led to desired or undesired outcomes can feed into future programs
- It is preferable to identify that a program is not achieving its desired outcomes before the program end, to give the potential of refining, re-defining or clarifying the program model. However, this requires considerable flexibility
- Process evaluation can inform stakeholders what is going on at ground level, as it is common for plans to change considerably on implementation
- Process evaluation can incorporate and cater for unexpected happenings or intermediate outcomes
- Process evaluation can build in flexibility and iteration into a program model from the onset
- Process evaluation, if done collaboratively, can lead to team building, as it can strengthen a common understanding of program aims, objectives and purpose.

3. PROGRAM EVALUATION

3.1. RELEVANCE OF PROGRAM EVALUATION AS A DISCIPLINE TO EXTENSION

The literature on evaluation is vast, constantly growing and changing in emphasis. There are now at least 17 international journals on evaluation and the authors have found over 5000 journal articles with program evaluation as a key subject. The discipline has its own societies, books, awards, conferences and standards. The field of evaluation today is characterised by great diversity: from large-scale, long-term, international comparative designs costing millions of dollars, to small-scale short evaluations of a single component in a local agency.

An historical perspective on evaluation can shed some light on the nature of evaluation as a discipline. Most of the early advances in program evaluation were in the field of education, and today the majority of literature still has an educational context. Second in magnitude is the context of health-related programs. Among the 700 members of the Australasian Evaluation Society, only two list agriculture as an interest in the 1996 membership booklet. When literature on the evaluation of agricultural programs is reviewed, it is immediately clear that limited information is available. In contrast, evaluation in education and health-related organisations is taken seriously and the field is constantly evolving to meet the needs of new programs and new ways of thinking. Evaluation of agricultural extension programs in the USA is also relatively active (Lees 1991). In general, the evaluation of agricultural extension programs in Australia appears to have been relatively stagnant, standard in method, and under-used.

It is important to review the advances that have been made in program evaluation globally. Much of the theory of evaluation was developed in the United States and it is relatively recently that program evaluation has come to Australia. A review of program evaluation in general should help to understand what has been tried and what has not been attempted in the context of agricultural extension in Australia.

3.2. DEFINITION OF PROGRAM EVALUATION

There is no single recognised definition of program evaluation. Historically, the definitions of evaluation and program evaluation have changed constantly, reflecting the evolving trends in thinking and practice. Early evaluation texts defined evaluation narrowly as 'the application of social science methods to measure goal attainment' or 'the systematic investigation of the merit or worth of an object' (Stufflebeam 1994). More recent definitions include the concept of using the evaluation findings for program improvement, development and decision making.

For the purposes of this study, program evaluation will be considered in its broadest sense, including 'quasi-evaluation forms' and the following definition is adopted:

"The process of providing information designed to assist decision-making about the object being evaluated." (Owen 1993)

3.3. DEVELOPMENT OF PROGRAM EVALUATION AS A DISCIPLINE

Program evaluation evolved primarily in the USA and became a semi-professional[1] discipline around 1960. The theory and practice of program evaluation that emerged at this time had roots in earlier work by Tyler (1967) in education, Lewin (1948) in social psychology and Lazarsfeld in sociology (Lazarsfeld and Rosenburg 1955). Program evaluation also has roots in the rapid economic growth in the US after World War II, in the interventionist role that the US federal government took on social policy during the 1960s and in the increasing interest in evaluation as an academic pursuit (Shadish *et al.* 1995).

At the end of the 19th century Joseph Rice conducted the first generally recognised formal education evaluation in the USA. The results of the evaluation led educators to revise their approach to the teaching of spelling. At the same stage the 'accreditation' evaluation movement began, even though it did not gain ground until the 1930s.

1900-1945 USA

During the early part of the 20th century a focus upon systemisation, standardisation and efficiency was seen in the field of education program evaluation. A large number of surveys were carried out in this period, which focused upon school/teacher efficiency. These surveys were externally evaluated. After World War I, school districts used these tests to make inferences about program effectiveness. Unlike later school curriculum evaluations, these surveys focused more upon local needs of the school district. In areas other than education the same focus upon systemisation was seen (Freeman 1977).

In the period between 1930 and 1945, the work of Ralph Tyler came to have an enormous effect upon the field of evaluation. Tyler coined the term 'educational evaluation', which meant assessing the extent to which valued objectives had been achieved in an institutional program. Evaluation was conceptualised by Tyler as a comparison of intended outcome with actual outcomes. It differed from the previously popular model of 'comparative experiment' in that it did not involve expensive and disruptive comparisons between experiment and control. Since Tyler's approach calls for the measurement of behaviourally defined objectives, it concentrates on learning outcomes instead of organisational and teaching inputs (Madaus *et al.* 1991). The current resurgence in outcomes-oriented evaluation, is clearly not a new concept for program evaluation!

In 1932 Tyler set up an eight-year study in which he evaluated 30 schools on the basis of objectives that were identified by the teachers themselves. This has been cited as the first comprehensive study of program evaluation (Gredler 1996). The concept of program appraisal was a new perspective introduced by Tyler.

[1] Program evaluation can be described as semi-professional as it shares certain attributes with other professions and differs from purely academic specialties. Program evaluation is not fully professionalised, like medicine or law; it has no licensure laws (Shadish *et al*. 1995).

1946-70 US

This period was characterised by rural poverty and despair in the inner cities in the US (Madaus *et al.* 1991). There was an expansion in educational services and social programs. The practice of standardised testing had expanded markedly. During this period evaluations were primarily the purview of local agencies. Federal and state agencies had yet to become involved. In 1957 the USSR launched the first space capsule (Sputnik) and the American public questioned their educational system; were American schools good enough to produce scientists and scholars to close the perceived technology gap? This led to a phase of centralised curriculum development and created the need for national curriculum evaluation.

'Great Society' programs in the US

Around 1960 the US government invested large sums of money towards programs in education, income maintenance, housing, and health. These programs are collectively termed 'Great Society' programs. Expenditure on these programs was massive and constituted a 1800 % increase in US public spending between 1950 and 1979 (Bell 1983). This vast expenditure raised issues which caused evaluation to be pushed forward as a necessary component of social programs: firstly, issues of accountability for distribution of funds; secondly, concern that program funds were being spent in ways that caused undesirable results. However, at that time of increase in demand for evaluation, there was a lack of personnel trained in this field.

The rest of the world

While large-scale programs with an evaluation component became commonplace in the US, this trend also occurred in Europe and Australia. By the late 1960's in the United States and internationally, evaluation research had become a growth industry (Wall Street Journal, cited by (Rossi *et al.* 1979).

1970 to early 1980s global professionalisation of program evaluation

Prior to 1970, limited published literature about program evaluation existed; many evaluations were carried out by untrained personnel, others by research methodologists who tried unsuccessfully to fit their methods to program evaluation (Guba and Lincoln 1981). Around 1973, program evaluation began to emerge as a semi-professional discipline. In the USA a number of journals including *Evaluation Review* and *New Directions in Program Evaluation* went to press. In the USA in 1970s, two professional evaluation societies were formed: the 'Evaluation Research Society' and the 'Evaluation Network'. In 1984 these two societies merged to become the 'American Evaluation Association'. By 1984 evaluation had become more international and the 'Australasian Evaluation Society' had formed along with the 'Canadian Evaluation Society', and by 1995 new professional evaluation societies had formed representing Central America, Europe and Great Britain. In 1981 a joint committee appointed by 12 professional organisations issued a comprehensive set of standards for judging evaluations of educational programs and materials.

Around this time evaluators realised that the techniques of evaluation must:

- serve the needs of the clients
- address the central value issues
- deal with situational realities
- meet the requirements of probity
- satisfy the needs for veracity. (Madaus *et al.* 1991).

Explosion of theories and techniques

Responding to the need for more sophisticated evaluation models, by the 1970s many new books and papers were published (Cronbach 1963, Scriven 1967, Stake 1967). Together these publications resulted in a number of new evaluation models described by the authors as responsive to the needs of specific types of evaluation. This body of information revealed sharp differences in philosophical and methodological preferences, but it also underscored a fact about which there was much agreement: evaluation is a multidimensional technical and political enterprise that requires both new conceptualisation and new insights into when and how existing methods from other fields might be used appropriately (Worthern *et al*.1997).

Despite all the new promising techniques, Madaus *et al.* (1991) point out that there is still a need to educate evaluators in the availability of new techniques, to try out and report the results of using these new techniques, and to develop additional methods. The emphasis must be on making the methodology fit the needs of the society, its institutions and its citizens, rather than the reverse (Kalpan 1994).

1980-1990s era of accountability

In the US a new direction for evaluation emerged in the 1980s with the presidency of Ronald Reagan who led a backlash against government programming. This is partly explained by the perceived failure of the earlier 'Great Society Programs'; social indicators revealed that many multi-million dollar projects had resulted in little social change. In the 1990s the call for greater accountability became central to programs at every level: national and local, NGOs and the private sector (Brizius and Campbell 1991). Indeed, in the late 1980s to 1990s there was growing despair about the fact that 'nothing seemed to work'. In 1988 Brandle's message in an address to professional evaluators was:

"No demonstrable relationship exits between program funding levels and impact, that is, between inputs and outputs; more money spent does not mean higher quality or greater results."

Brandle went on to say that "until systematic accountability is built into government, no management improvements will do the job".

3.4. THE DEVELOPMENT OF EVALUATION IN AUSTRALIA

Although the root of evaluation development lies in the US, in the 1960s evaluation began to surface in Australia and later in Europe. The Australasian Evaluation Society was the first non-American evaluation society to emerge. It has been remarked that the development of program evaluation in Australia differed somewhat from that in the US, in that internal evaluation was always more prevalent than in the US. This can perhaps be explained by the attitude in the US to evaluation as an academic discipline and the existence of the numerous evaluations higher degree courses there. In the US, evaluation tends to be seen as something that is done by trained external evaluator, qualified with a PhD, whereas in Australia there tends to be a greater focus on internal evaluation (Winston J. 1997 per. comm).

4. FORMS OF PROGRAM EVALUATION

4.1. OWEN'S FIVE FORMS OF EVALUATION

Various attempts have been made to classify evaluation *per se*, some by categorising forms of evaluation by purpose (Owen 1993), others by methodology (Stake 1973), and others by the position of the major audience (Worthen *et al.* 1997). In the Review, the conceptual meta-model of Owen (1993) is adopted, to classify examples of evaluation studies obtained from the field. This meta-model uses a relatively straightforward framework for classifying evaluation approaches into five categories or forms, based on purpose:

- Form 1 evaluation for impact assessment
- Form 2 evaluation for program management
- Form 3 process evaluation
- Form 4 evaluation for design clarification
- Form 5 evaluation for program development.

Owen's meta-model is further elaborated in Table 1. Although in practice, many evaluation studies span more than one of Owen's forms, the framework is a useful tool in conceptualising evaluation *per se*. The following sections will expand on each of these forms and provide case study examples of their application in the context of agricultural extension. It should be pointed out at this stage that not all theorists agree that the last two forms (evaluation for design clarification and evaluation for program development) are evaluation at all. Case studies of Form 5, evaluation for program development, are only briefly touched upon in the Review as in agricultural extension this is largely called needs (or situational) analysis rather than evaluation. There is good reason, however, to include a description of Form 5 in the Review, as a thorough approach to developing programs in Australia could have marked benefits in terms of improving the effectiveness of agricultural extension.

Table 1 Owen's five forms of program evaluation

Form or approach of evaluation	**Form 1** Evaluation for impact assessment	**Form 2** Evaluation for program management (monitoring)	**Form 3** Process evaluation	**Form 4** Evaluation for design clarification	**Form 5** Evaluation for program development
Ultimate Purpose	Justification	Accountability	Improvement	Clarification	Synthesis
State of program	The program is settled	The program is settled	The program is at developmental stage	The program is at developmental stage	The program has not yet been implemented and is at planning stage
Timing of evaluation	After program	Can run throughout the program	During program implementation	During program implementation	Before program is implemented
Focus of evaluation	On outcomes /delivery	On delivery / outcomes of the program	On delivery of the program	On the design of the program	On context of the program environment
Typical approaches	Objectives-based outcome evaluation, needs-based evaluations, economic assessment of cost–benefit of program	Program monitoring, component evaluation, system evaluation, use of performance indicators over time	Implementation studies, action research, responsive evaluation, topical RRAs or PRAs, process evaluation.	Evaluability assessment, program logic, accreditation	Needs assessment, review of best practice, research synthesis, exploratory RRA (rapid rural appraisal) or PRA (participatory rural appraisal) ex-anti evaluation
Time span	Program finish	Throughout program	Program is settled	Program is developing	Program synthesis
Form 1	←	→			
Form 2	←				→
Form 3			←		→
Form 4				←	→
Form 5					← →

Adapted from (Owen 1993)

4.2. ALTERNATIVE META-MODELS FOR CLASSIFYING EVALUATION

Although Owen's (1993) meta-model is saluted as a relatively simple means of categorising forms of evaluation, several other classifications are reported here from the literature. According to Smith (1994), classifications provide a structure for selecting alternatives and identifying gaps to be filled by new approaches. The classifications employ different dimensions of contrast, such as purpose, nature of question addressed, typical methods, developer and proponents, key elements and major audiences. In reality these classifications are over-simplistic because of the complexity of the models that they attempt to classify. Wadsworth (1997) points out that "One would really need a 3-dimensional map on which to try to plot a full picture and even then some of the techniques would have to be moved back and forth between one category and another".

Despite the fact that these meta-models can present difficulties when used to classify real evaluation examples, they can provide an effective conceptual framework to help understand the different dimensions of the complex world of program evaluation. Some examples of different meta-models of evaluation are listed below:

- Stake's meta-model (1973) appears to use criteria primarily associated with the **methods** of evaluation used

- House (1978) uses a combination of typical **methods** and **purpose** of the evaluation

- Stufflebeam and Webster (1981) present a classification that appears to be based upon **the position of the major audience, and the methods** employed in the evaluation

- Worthen *et al.* (1997) present a meta-model that appears to be based upon the **position of the major audience**

Smith (1994) classifies evaluation in terms of seven **approaches** that are referred to as distinct from **models,** in that they are, 'broader, less focused conceptualisations that include grouping several related models'. His classification appears to be based on the previous classifications that were carried out by Stake (1973), House (1978), Stufflebeam and Webster (1981) and Worthen *et al.* (1997).

4.3. THE DOMAINS OF EVALUATION

The Review is organised on the basis of 'forms' of evaluation and discusses various frameworks that can be employed and their associated methods. Much of the literature in program evaluation discusses various 'models' of evaluation. These various 'domains' in evaluation (form, framework, model, and method) can be confusing, so an attempt is made here to explain the relationships between various terms in common use in evaluation literature. Figure 3 illustrates the conceptual relationship between the different domains of evaluation for a hypothetical program.

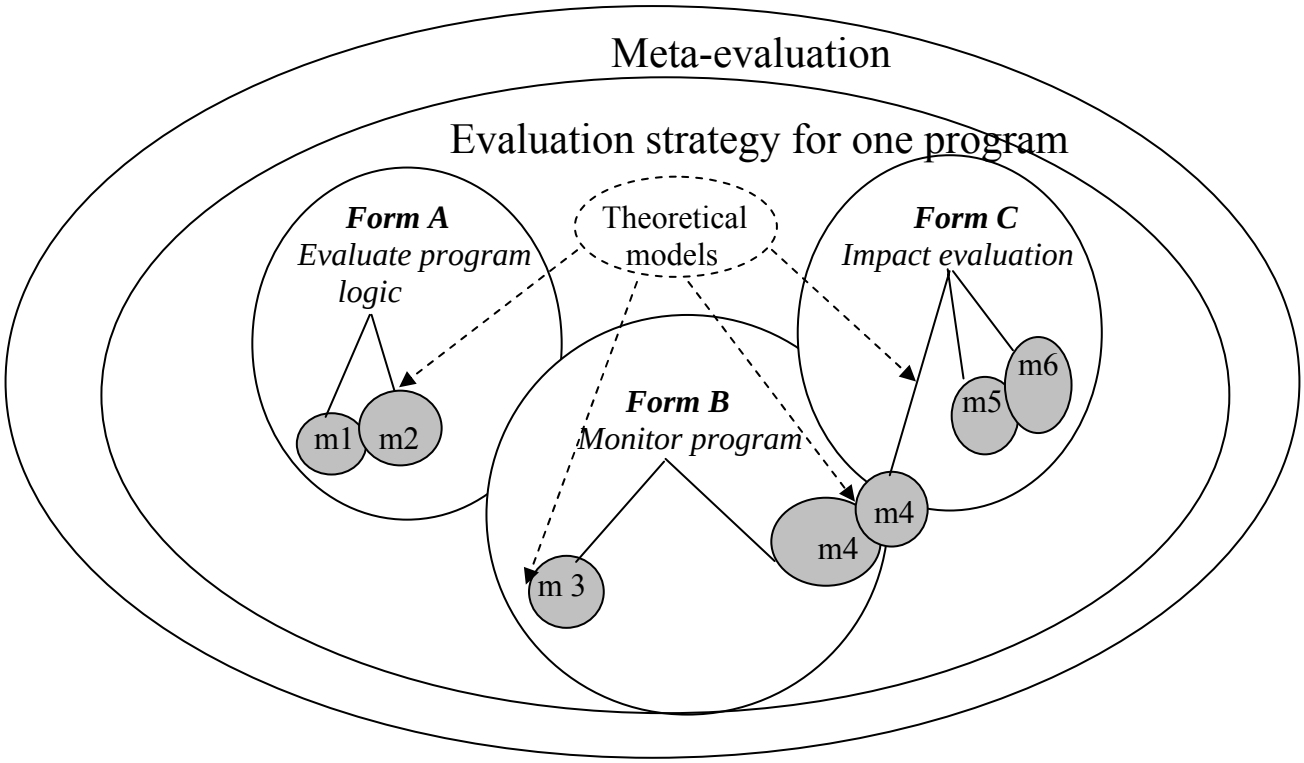


Figure 3 Domains of evaluation for a hypothetical program

In Figure 3, the outer oval represents the domain of 'meta-evaluation', where various program evaluations are evaluated themselves. The inner large oval represents the domain of evaluation strategy for an individual program. Within this oval are egg shapes representing different forms of evaluation. For this hypothetical program, three forms of evaluation are required to meet three different evaluation purposes. Conceptual models (dashed lines) influence the choice of methods (grey circles). It should be noted that one method of evaluation could be used for more than one form of evaluation.

Methods and models have been the focus of many publications on evaluation. One example of a model of evaluation is 'goal-based' evaluation. This work does not attempt to review models or methods of evaluation although these are mentioned in summary form in Section 13 and where included in the case studies in Sections 7-11. For this the reader is referred to texts such as *Evaluation Models* by Madaus *et al.* (1991) or *Foundations of Program Evaluation* by Shadish *et al*. (1995). The Review takes the view that no one model or method of evaluation is appropriate in all evaluative contexts. For this reason we have concentrated on the various forms of evaluation, which are a function of purpose. In the Review we differentiate between the terms:

- meta-model of evaluation
- framework for evaluation strategy
- forms of evaluation
- conceptual models of evaluation
- methods of evaluation.

In many texts these terms may overlap. In the Review the terms are given more specific meanings, which are presented below.

Meta-model is a framework for describing or categorising various different approaches or forms of evaluation. An example of this is Owen's meta-model in which he describes five different forms of evaluation that occur in program evaluation today.

Strategic frameworks for evaluation. An organisation, program or project, may conduct several different forms of evaluation to meet various needs. For example, these forms may include on-going monitoring systems to feedback information on progress, one-off evaluations to investigate a particular element of the program, and a final impact assessment to determine the extent to which the goals of the program have been reached. These forms of evaluation together make up the evaluation strategy that is illustrated as the inner oval in Figure 3. Section 12 illustrates some strategic frameworks for evaluation in the field of agricultural extension.

Forms of evaluation. In the Review we have taken **forms** of evaluation to be a function of purpose. One form of evaluation might aim to understand whether the needs of the beneficiaries were being met. Another form of evaluation may judge whether the predetermined outcomes have been reached. In Sections 7-11 five forms of evaluation are described, using case studies to illustrate each form.

Model is used with wide variability and considerable ambiguity in evaluation practice and literature. Generally it refers to particular conceptions, approaches, methods and even loose theories for thinking about and/or conducting evaluations (Smith 1994). In the Review it is used to signify conceptual ideas based on particular paradigms of how an evaluation should be conducted. These models may be implicit or explicit: implicit in that they are carried around in our heads and based on our common sense; explicit in that they follow a particular conceptual model of how evaluation should be carried out. Examples of models include utilisation-focused evaluation, goal-free evaluation and naturalistic inquiry.

Methods of evaluation are the tools of the trade. One particular method may only be appropriate for a particular context. Here, we have taken the view that evaluation strategy and form of evaluation should be considered prior to selecting methods. There are potentially thousands of methods, including mail survey, cost–benefit analysis, interview, questionnaire, and goal attainment scaling. In Figure 3 the methods are illustrated by grey circles and each form of evaluation may use several methods. A summary of methods may be found in Section 12.

4.4. DIFFERENT LEVELS OF PROGRAM EVALUATION

Programs that are implemented at different levels of the planning hierarchy usually require different approaches and methods of evaluation. The Review is largely concerned with the evaluation of programs (or projects). Another level of evaluation is that of 'one-off events' such as meetings, field days or training courses. Approaches and methods for the evaluation of 'one-off' events are dealt with in texts such as that of Wadsworth (1997). For information on the evaluation of policy, readers are referred to Weiss and Bucuvalas (1977).

5. CHARACTERISTICS OF EVALUATION IN AGRICULTURAL EXTENSION

5.1. AIM AND METHOD

In the Review, documents were collected from over 80 evaluation studies from a wide range of sources, ranging from rural extension centres, to universities and agriculture departments in various states, to private consultants and individuals in the field of agricultural extension. These documents came from all six states of Australia and were conducted within the past ten years. This list grew to over 100 respondents from all states and some overseas specialists, through advice and further contacts received from individuals and members of the Australasian Pacific Extension Network.

Of these documents, 50 were selected to provide the data for an analysis of evaluation in extension, which forms part of the Review (listed in Appendix 3). The remaining reports were excluded as they contained insufficient information, were primarily evaluations of 'one-off' workshops, were guides as to how to conduct a hypothetical evaluation, or were received too late. The main method of presenting the results of the analysis was to sort the sample of evaluation studies according to Owen's (1993) classification of evaluation forms and to provide case studies to represent each form. Attempts were also made to analyse the information gathered in ways that may provide insights into the range of evaluation studies being conducted within the field of agricultural extension.

However, the sample of 50 reports studied may not be fully representative of the true range of evaluation in agricultural extension being conducted. Many evaluations are conducted internally or informally by departments and not written up, while others are documented but the results remain confidential.

The sample was obtained from evaluators active in the field who were asked to provide materials and also the names of other people involved in evaluation. The initial list of respondents included all those professionals in the Institute of Agricultural Science who listed evaluation as an interest. A very common response to our request for information was that practitioners were just beginning to learn about evaluation and had no written reports that they felt were good examples of evaluation to include as case studies in the Review. This result confirms the findings of Woods *et al.* (1993) and Russell *et al.* (1989) in their reviews of the state of extension practice in Australia.

5.2. CRITERIA USED TO CHARACTERISE THE RANGE OF EVALUATION STUDIES

The 50 evaluation studies in the sample were characterised according to the following variables relating to (a) the program, (b) the evaluation, and (c) the type of methods used:

About the program

- level of program ie evaluation of program at regional, state or national level
- stage of program development – developmental, settled or completed
- field or aim of program – education and skill development, community development, environmental conservation, or transfer of technology
- level of participation of program – using an index of participation (Appendix 1).

About the evaluation

- evaluation approach – closest fit to Owen's forms and whether the study was
- primarily formative, summative or both
- broad nature of methods – primarily qualitative, quantitative or mixed
- level of client participation in the evaluation – using an index (Appendix 2)
- evaluation an internal process, or carried out by external consultants
- evaluation commissioned by insiders or outsiders
- levels of Bennett's hierarchy considered in the evaluation
- needs-based or goal-based evaluation.

Methods used in evaluation

- single or multi-perspective view maintained
- statistical inferences made
- use of direct quotations from stakeholders
- use of structured questionnaire
- use of focus-group discussions
- use of rapid rural appraisal
- application of experimental design
- medium of inquiry used – telephone, face-to-face interview, or postal survey.

5.3. SOME RESULTS AND DISCUSSION

A summary of the findings of the analysis of the 50 evaluation studies in the sample is provided in Table 2.

Table 2 Summary of main characteristics of the sample of evaluation studies

Criterion	Approximate percentage in categories (n=50)										
	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Form	summative 54%						formative 32%				both 14%
Stage of Program	developing 25%				Settled 58%						finished 12%
Perspectives considered	multi-perspective view taken 40%						maintained a single focus-perspective 60%				
Evaluator	external evaluator 63%							internal evaluator 26%			both 11%
Audience	external audience 20%				Internal audience 65%						Both 15%
Data Type	qualitative 44%						quantitative 27%				both 11%
Surveys	structured survey 52%						no structured survey 48%				
Focused group discussion	focused groups used 42%					not used 56%					

The form of evaluation

When studies were classified by means of Owen's five forms (Owen 1993) a difficulty experienced was that studies could often be placed in more than one category, especially with regard to the time frame that Owen attached to the meta-model (see Table 1, Section 4). The distribution of the studies according to a slightly modified Owen categorisation is shown in Figure 4.

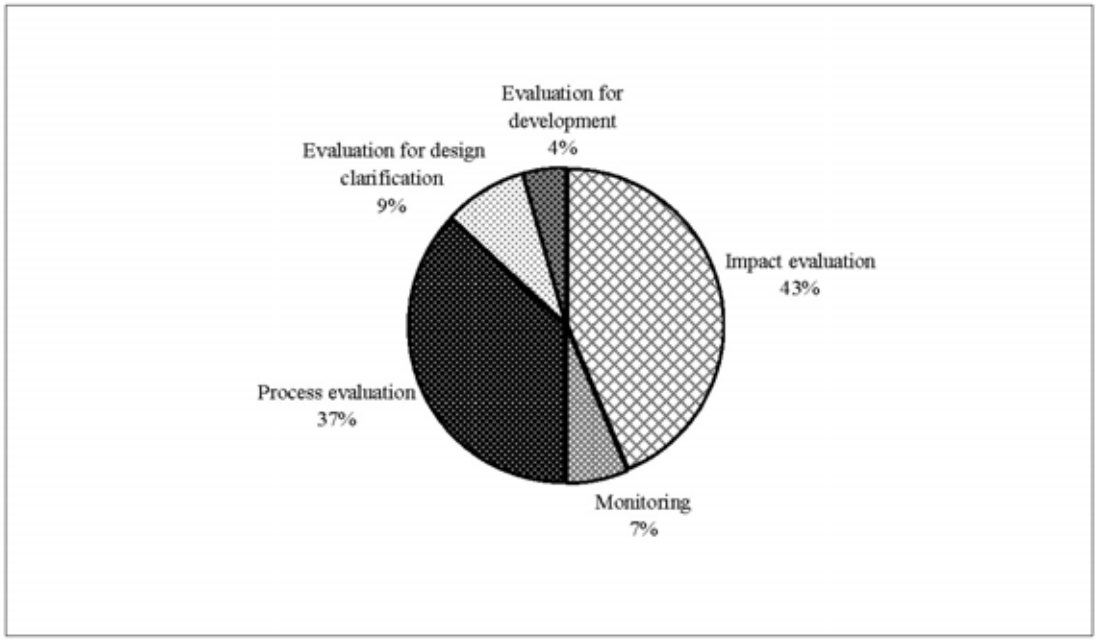
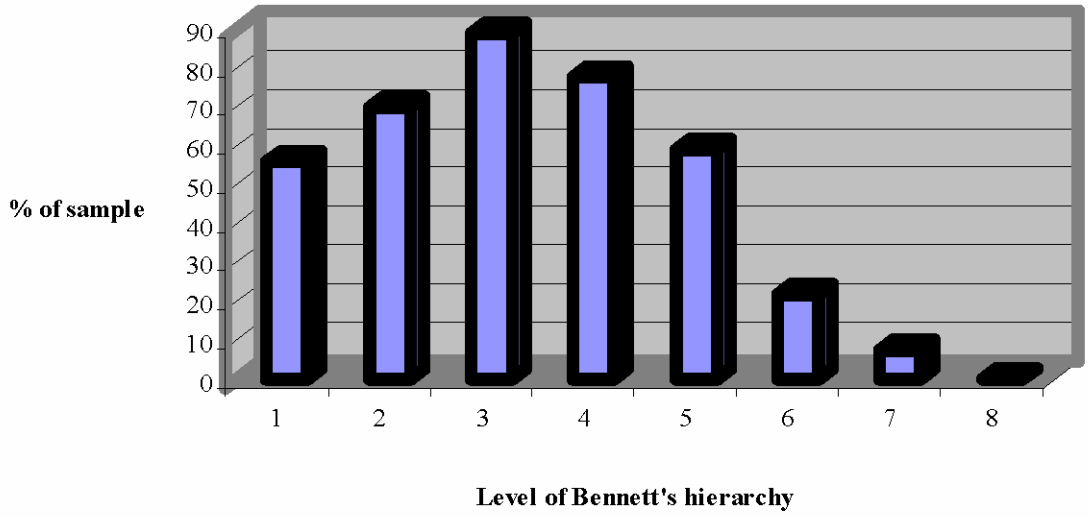


Figure 4 Distribution of the evaluation studies according to Owen's five forms

The distribution in Figure 4 shows that impact studies were the most common, but that process evaluation (Form 3) was also very common. From information on the timing of the studies there appears to be a recent trend towards process evaluation, and also towards the use of evaluation for formative rather than summative purposes in agricultural extension.

Analysis of the sample using Bennett's hierarchy

Bennett's hierarchy is a hierarchy of goals established for agricultural extension programs (Bennett 1975). This framework is generally well-known to agricultural extension agents, probably because it is mentioned in a well-used extension text (Van den Ban and Hawkins 1996, 205-208). Bennett lists seven levels of goals in extension, and claims that it becomes more difficult to evaluate at higher levels of the hierarchy, as it becomes more difficult to show that changes at these levels are the result of extension activity and not of other factors. An eighth level (consequences for society) is an additional level that appears in the Van den Ban and Hawkins text (1996).



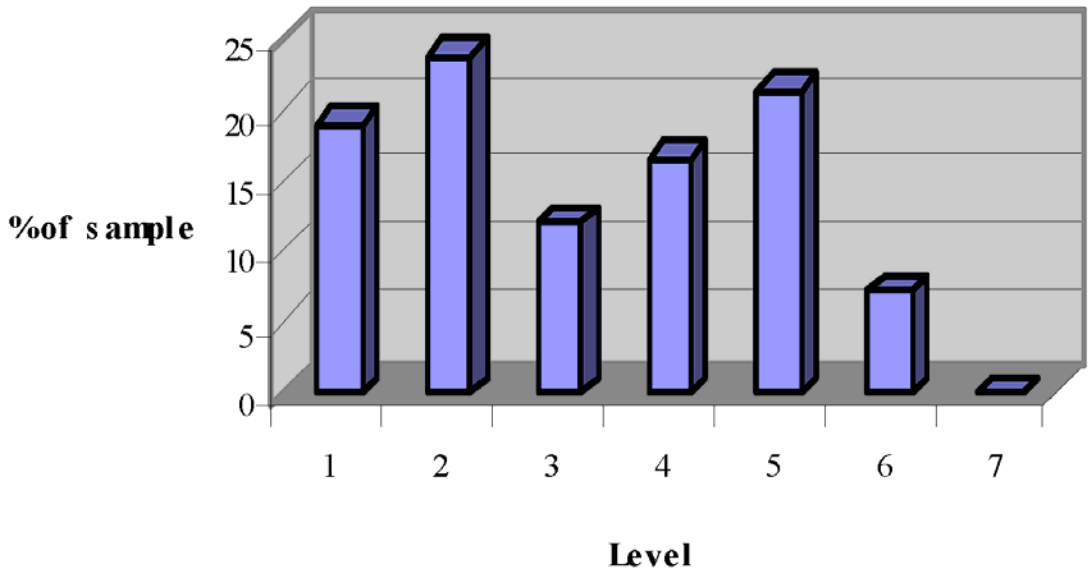
Level 8	Consequences for society (an additional level to Bennett's hierarchy)
Level 7	Consequences for the target group
Level 6	Behavioural changes in the target group
Level 5	Change in knowledge, attitude, skills, motivation and group norms (KASA)
Level 4	The farmer's opinion about extension activities
Level 3	Farmer participation in extension activities
Level 2	Implementation of the program by extension agents
Level 1	Programming of the extension activities

Figure 5 Distribution of evaluation studies according to the level of investigation according to Bennett's hierarchy

Figure 5 shows the frequency of occurrence of studies in the sample at each level of Bennett's hierarchy. The majority (89%) of the evaluations considered level 3; the degree of farmer participation in farmer activities, while 60% of the evaluations considered changes in KASA (level 5), and only 22% measured behavioural changes. Obviously it is difficult to evaluate at levels 7 and 8, especially when it is considered how many different
actors are involved in the agricultural information system. It would have been interesting to study the change in orientation of evaluations of extension programs over the last 30 years, but most of the evaluations studied were fairly recent.

Level of participation by clients in the evaluation

Assessing the level of participation of clients in a program (or an evaluation) can be very subjective unless well-defined criteria are used as a means of scoring participation. Figure 6 shows the results of scoring the evaluation studies in the sample against a continuum of participation (see Appendix 2) based on Arnstein's (1969) participation scale. In 54% of the evaluations the clients were not heavily involved in developing the indicators for the evaluation. At the other extreme, 7.5% of the evaluations involved the clients being 'facilitated' to carry out their own evaluation.



Participation levels adopted (see Appendix 2)						
1	**2**	**3**	**4**	**5**	**6**	**7**
External evaluator, no consultation	External evaluator, staff consulted	Staff conduct evaluation, clients not consulted	External evaluation, staff & clients consulted	Staff & clients develop indicators together	Clients conduct facilitated self-evaluation	Clients conduct self-evaluation - alone

Figure 6 Level of participation in evaluation by farmer clients in the studies

In many instances in agriculture, evaluators have seen little reason for client participation. However, the proponents of highly participatory evaluation stress the relevance in many situations in agricultural extension of involving clients in developing the evaluation framework, especially with regard to the choice of indicators. Patton (1997) writes that such 'empowerment evaluation' is most appropriate where the goals of the program include helping participants become more self-sufficient and personally effective. In such instances, empowerment evaluation is also 'intervention oriented' in that the evaluation is designed and implemented to support and enhance the program's desired outcomes.

Most common approach to evaluation in agricultural extension

The wide range of studies encountered in the Review indicate that it would be unreasonable to attempt to describe the typical evaluation in agricultural extension in Australia. However, after reading over 100 reports of evaluations, the authors' impressions are that the most common approach to evaluation nowadays in agricultural extension in Australia is a summative evaluation, using both qualitative and quantitative data. Evaluations were usually conducted while the program was in a settled stage of development and carried out by external evaluators. The clients of the program (usually the farmers) would generally not be involved in developing the indicators for the evaluation, although a pre-test of surveys would be carried to ensure that questions were well phrased and understandable. The general aim would be to provide a report to justify spending and to understand whether the stated objectives of the program had been met. A focused-group discussion might be used with a small number of farmers and extension agents to determine the factors most important to them.

This study did not set out to examine changes in the type or approaches to evaluation. However, there are strong indications from the sample of reports received that the range of approaches to evaluation existing in agricultural extension is expanding. There are a number of new and innovative evaluations under way that are generally more formative and qualitative than appears to have been the norm in the past.

Expertise and confidence of extension agents in evaluation

Anecdotal evidence from discussions with various organisations and individuals during this study suggested that there was a lack of confidence as to how to select appropriate approaches and evaluation tools, conduct evaluations and utilise findings. The observations suggest that in agricultural extension there is a tendency to stick to a handful of familiar frameworks and methods such as Bennett's hierarchy (Bennett 1975). In some of the studies there seems to be a lack of understanding of evaluation in its broad sense, and as a result practitioners confine themselves to familiar methods that are not appropriate in all situations.

On the other hand, some of the evaluation documents reviewed revealed studies that were very innovative and exploratory in approach. These examples generally fit under the umbrella term of 'process evaluation'. This work is largely developing from the school of adult learning, or action research, although these approaches are not always necessarily labelled as evaluation.

Chen (1990) in his book *Theory-driven evaluation* discusses the traditional neglect of program theory and the popularity of method-oriented evaluation. He writes:

> 'Until very recently evaluation literature has rarely been concerned with the importance of theory in evaluating a program or with how to incorporate theory into evaluation processes'.

The results of the Review of evaluation in the field of agricultural extension concur with Chen's findings, although there is indication of a move away from method-driven evaluation in extension. This movement appears to be occurring where agricultural extension programs are incorporating elements of adult learning theory into their programs.

An important factor contributing to a lack of expertise and confidence in evaluation is the fact that most extension personnel are trained primarily in the physical sciences and few have a sound background in social science methods. Nowadays extension agents are expected, on top of the requirements of applied science-based aspects of their work, to cope with a range of concepts from the social sciences. Most have never studied literature on program evaluation, and if they have had training in social science methods it has been mainly from in-service training on short courses. There is an obvious need to address this problem if extension organisations are to improve the ability of staff and the quality of evaluation practice.

6. CASE STUDIES OF DIFFERENT FORMS OF EVALUATION IN AGRICULTURAL EXTENSION

In the following Sections (7-11), five forms of evaluation adapted from Owen's meta-model (Owen 1993) are outlined and each of the forms are illustrated with case studies from agricultural extension. The focus of these case studies is more upon the form of evaluation than on methods or recommendations or findings. Modern texts on evaluation emphasise that for effective, useful evaluations, the purpose of the evaluation must be made very clear first, and that only when all the stakeholders achieve a consensus on why the evaluation is being carried out should the methods be considered.

So the methods of evaluation described in the case studies selected are incidental. A particular method might or might not be appropriate for use across a range of forms of evaluation. For instance, 'cost–benefit analysis' is a well-known method that has been used on its own or with other methods across a number of forms of evaluation, but is not a form in itself (see Section 7.3). A list of methods used in the evaluation of extension is provided in Section 13.

A summary table is presented in the box above each case study, to give the reader a rapid impression of the type of evaluation being presented, an example is given in Table 3. The position of the evaluation with regard to a continuum of 'soft' to 'hard' evaluation is indicated first by the position of the bullet point in the continuum. In this context an extremely 'soft evaluation' refers to a study carried out using qualitative data, which is aimed at improving a program, carried out by internal evaluators, formative in nature, maintains a multi-perspective view and which is informal. An extremely 'hard' evaluation would be a quantitative evaluation carried out by external evaluators, oriented towards justification of program spending rather than improvement of an existing program, and involving a more formal approach aimed at an external audience.

Table 3 shows an example of a summary table relating to a case study that describes an evaluation that was at the 'hard' end of the continuum. The position of the black dots indicate that: it was a summative evaluation, it had most in common with Owen's Form 1 (impact evaluation), it used mainly quantitative data, it was conducted when the program had already finished and the program had been planned at a national level.

Table 3 An example of the summary table preceding each case study

Evaluation	*Soft* ◄——————————●——————————► *Hard*				
Form	*Formative*			*Summative* ●	
Owen's form	*F5*	*F4*	*F3*	*F2*	*F1* ●
Data collection	*Qualitative*		*Both*		*Quantitative* ●
Stage of program	*Forming*	*Developing*		*Settled*	*Finished* ●
Program planning level	*Regional*		*State*		*National* ●

7. FORM 1: EVALUATION FOR IMPACT ASSESSMENT

Owen describes 'impact evaluation' as evaluation that leads to a decision about the worth of a program and which often has a strong summative emphasis. This type of evaluation is generally carried out at the end of the program, or when a program is at a settled phase. The principal foci of impact evaluations are to:

- understand the outcomes of the program
- justify program spending
- gain guidance about what to do next.

Within the category of impact evaluation in agricultural extension, several approaches were found. Each of the following approaches is illustrated with a case study from extension:

- goal-based impact evaluation
- needs-based impact evaluation
- comparative economic impact evaluation
- impact evaluation for illumination.

A distinction can be made between '**goal-based'** evaluation and '**needs-based**' models of evaluation. These two models are dealt with separately below, with case studies provided for each. Another form of impact evaluation is based on financial analysis of costs and benefits and an example is given under Section 7.3. Cost–benefit analysis can also be used to compare the impact of alternative programs. A fourth example of impact evaluation (shown under Section 7.4) is evaluation for illumination – which aims to explore both the intended and unintended outcomes of the program interventions.

7.1. GOAL-BASED IMPACT EVALUATION

Goal (or objectives) -based evaluations judge whether the stated goals or objectives of a program have been achieved. Tyler (1967) was amongst the first to develop and use goals-based evaluation. In this approach goals are taken as given, and decisions about the success of a program are deduced from whether the program has met its goals. According to Owen (1993) the main tasks in planning a goal-based evaluation are to:

- determine the 'real' goals of the program
- decide how to determine whether the program has led to the attainment its of goals.

A standard procedure in goal-based evaluation is to develop measures of goals that have strong face validity. However, sometimes it is necessary to use substitute measures or indexes that stand in the place of preferred or ideal measures. An example of a goal-based evaluation is given in Case Study 1.

Case Study 1 Landcare: does it make a difference?

Evaluation	*Soft* ◄——————●——————► *Hard*				
Form	*Formative*			*Summative* ●	
Owen's form	*F5*	*F4*	*F3*	*F2*	*F1* ●
Data collection	*Qualitative*		*Both*		*Quantitative* ●
Stage of program	*Forming*	*Developing*		*Settled* ●	*Finished*
Program planning level	*Regional* ●		*State*		*National*

The program

The Landcare program emerged in the state of Victoria during 1986 and has been spread throughout Australia as a model for effective community action to manage land degradation and assist the move to more sustainable land use. The evaluation described here was carried out in 1993, at a time when the Landcare program in Victoria could be described as in a settled state, with many groups having been operating for 6 years. This evaluation forms part of a wider Landcare evaluation undertaken within the Johnstone Centre in Wagga Wagga. This program is based on the assumption that Landcare group actions will facilitate the process of community development and thereby produce more aware, informed, skilled and adaptive resource managers with a stronger stewardship/land ethic, and result in the adoption of more sustainable natural resource management practices.

Goal of the evaluation

To assess the effectiveness of Landcare groups by investigating the work of groups and their impact upon participant's behaviour and other rural landholders. Specifically, the study aimed to determine whether Landcare participants and non-participants were significantly different in terms of key social and farming variables and whether Landcare participation made a significant difference to landholders' awareness of issues, level of knowledge of key resource management, and their stewardship ethic.

Form of evaluation

The overall purpose was to investigate some aspects of the overall program aims. Owen's Form 1, 'Impact Evaluation', best covers this example. This evaluation made summative judgements as to whether the program had met intermediate goals of the Landcare program, so took a goal-based approach. This evaluation attempted to measure the degree to which the goals (or substitute goals) have been met by program activities.

Method of evaluation

A regional case study was carried out. The study comprised a comparison between Landcare respondents with non-Landcare respondents, and between respondents in Landcare areas with respondents in areas where Landcare groups did not exist. In April 1993 a 16 page survey was mailed to all land managers of rural properties in 12 catchment areas (9 Landcare areas and 3 non-Landcare areas) of North Victoria. Surveys were addressed to one listed owner per rural property, 30% of who were women. The researcher's knowledge of the area and links with Landcare group there played a critical factor in the selection of the region. All groups surveyed had been operating for more than 2 years. The surveys were pre-tested using 2 focus-group discussions (FGD) of Landcare participants and non-participants at separate locations. The individuals were contacted by telephone and given a brief overview of the project and invited to the pre-test FGD. The pre-test participants were mailed a survey before the FGD took place. Important revisions of the survey were made as a result of recommendations following the discussions at both meetings.

The questionnaire was a small (B5) booklet with a distinctive, authoritative cover and was posted with a covering letter and stamped return envelope. A reminder / thank you letter was posted after 8-10 days and a second mailout occurred after 6 weeks. Considerable effort appears to have been put into selecting appropriate topics to include in the survey. Factors such as property size, hours worked off-farm, length of experience as a farmer, membership of other groups, age of respondent and school education were correlated against participation in Landcare.

Use of findings
The results were quantitatively analysed using chi-squared and multivariate analysis. The statistical findings were then represented in tables with accompanying text to explain the findings. These findings were presented and published in a booklet.

Source: Summarised from a report by Allan Curtis and Terry De Lacy (1994)

7.2. NEEDS-BASED IMPACT EVALUATION

Scriven (1967) offered a different perspective towards impact evaluation. He suggested that in some circumstances goal-based evaluation has limitations and introduced the concept of '**goal-free evaluation**'. His argument was that testing the goal of a program does not tell the worth of that program for society in general, in that *it does not assess the goals themselves*.

A goal-free evaluation in the strictest sense would occur without the evaluator being aware of the stated goals of the program. Critics of the goal-free approach to evaluation state that in a goal-free approach it is the evaluator who decides the needs of society, and that the evaluators' goals are substituted for those of the project (Patton 1997). In reality, goal-free evaluation in its strict sense is rarely practiced in program evaluation. In this text the term '**needs-based**' evaluation will be used to refer to evaluation where the orientation of the evaluation is directed toward the needs of society or the stakeholders rather than a judgment of whether the given goals were reached. For example, if the goal of a hypothetical program were to produce a million tonnes of beef as cheaply as possible, a strictly goal-based program would work out whether this goal (and perhaps intermediate goals) had been achieved. A **needs-based** evaluation, in addition to investigating whether the given project goals have been met, might also look at whether a million tonnes of beef was needed, what impact this increase in production would have on the environment, and so forth.

The choice between a 'needs-based' approach to evaluation and a 'goals-based' approach should be made with regard to the underlying purpose of the evaluation. Patton (1997) suggests that to be an effective evaluator one needs to be able to evaluate with or without goals. The following Case Study 2 of the West Hume Landcare Group gives an example of a needs-based, impact evaluation. It is also an interesting case study as it uses very mixed methods.

Case Study 2

A case study undertaken with the West Hume Landcare Group as part of an international research project on Participatory Watershed Development

Evaluation	*Soft* ←●——●——→ *Hard*				
Form	*Formative* ●		*Summative* ●		
Owen's form	*F5*	*F4*	*F3*	*F2*	*F1* ●
Stage of program	*Forming*	*Developing*	*Settled* ●	*Finished*	
Data collection	*Qualitative*	*Both* ●	*Quantitative*		
Program planning level	*Regional* ●	*State*	*National*		

The program

West Hume Landcare group was chosen to be one of two case-study evaluations of the National Landcare program. The group was established in 1989. There are 160 farms in the area and 85 families are group members. The main land degradation issues are rising water tables and salinity associated with a massive change in the natural vegetation, soil acidity, soil structural and nutrient decline. The evaluation was carried out when the group was well established. Nationally, there has been growing concern about how to evaluate local Landcare activities in a way that is useful to local groups, but which can also provide information to assist broader-scale policy development and help give a national perspective on Landcare achievements. This concern led to LWRRDC providing additional funding for a case study and to support the active involvement of a National Landcare facilitator.

Goal of the evaluation

The overall goal of the evaluation was to appraise the success of Landcare, shortfalls and future opportunities. More specifically the goals were to:

- *enable Landcare group members to review the achievements, shortfalls and future opportunities for the group*
- *provide analysis of implications for Landcare policy*
- *understand the economic outcomes of the group effort*
- *provide an evaluation case study on Landcare to be presented at an international conference*
- *provide a draft manual report for LWRRDC on how other groups could undertake a study as this, using a PRA (participatory rural appraisal) approach.*

Form of evaluation

*This evaluation approach is '**needs-based**' as opposed to 'goals-based' in that it attempts to evaluate against the needs of the community rather than against predefined goals. It is most similar to Owen's Form 1, impact evaluation. However, it has both formative and summative elements: while it judges the performance of the project against community needs (summative), it also offers recommendations for the future (formative). This evaluation uses mixed methods ranging from participatory self-evaluation to economic analysis by external consultants.*

Actors

External evaluators were used but their role was that of facilitation and analysis. There was a high level of community involvement in the evaluation. The audience was both internal (group and Landcare program) and external (IIED, LWRRDS).

Methods of evaluation

4 methodological approaches were used:

1. *Qualitative data were gathered by means of a PRA. PRA is a method for rapidly gathering key information about a rural situation by drawing on local knowledge and integrating it with the knowledge and experience of professionals from a range of different disciplinary areas. The PRA was conducted for a period of 7 days in 1994. Team members came from a variety of disciplinary*

areas. PRA techniques included semi-structured interviews, group histories, questionnaires, preliminary analysis of cost–benefit of some common land conservation practices, and an economic study of 2 local farms and 2 hypothetical best management practice farms on whether farmers could afford Landcare.

2. *Some quantitative data were collected using a structured questionnaire, also during the PRA exercise, using semi-structured interviews.*

3. *Secondary information in the form of reports and previous studies were also examined and incorporated into the exercise.*

4. *Economic evaluation was carried out by external consultants prior to the PRA exercise. The purpose of that study was to determine whether farmers could afford Landcare. A detailed financial analysis was conducted of 4 typical farms of different sizes and enterprise mixes, and two hypothetical 'best practice' farms. Data for the analysis of the 4 farms were obtained by interview with 4 selected landholders.*

PRA: during the day, team members conducted focused group discussions with a range of stakeholder groups. In addition, members travelled to farms and interviewed individual farm families. Groups included Landcare members, non-members, senior members of the community, a Landcare committee, and a women's' group. Each evening the team presented and analysed the findings during group discussions. At the end of the week the first results of the analysis were presented to the community and team during a dinner.

Use of findings
As well as being written as a report, this case study was presented at an international workshop on participatory watershed development in Bangalore, India, amongst 23 case studies from different countries. The West Hume case study was presented by farmers and Landcare group members and was especially appreciated.

Source: Jim Woodhill, Paul Trevethan, Judy Frankenburg, Celine Beaulieu (1994), The West Hume Landcare group

7.3. FINANCIAL ANALYSIS OF OUTCOME

Cost–benefit analysis (CBA) is a procedure for comparing alternative courses of action (or no action) by reference to the net social benefits that they produce. A net social benefit refers to the difference between social benefits and social costs. CBA is a method for organising information to aid decisions about the allocation of financial resources (DOF 1991).

CBA can be applied to specific projects or to larger programs as a whole. Its power as an analytical tool rests in two main features:

- costs and benefits are each expressed (as far as possible) in money terms and hence are directly comparable
- costs and benefits are valued in terms of the claims they make and the gains they provide to the economy as a whole, so the perspective is a global one rather than that of any particular individual, organisation or group.

CBA is widely used as a generic tool in evaluating the financial impact of agricultural programs. Its application may be *ex-ante* (i.e. before implementation) to help:

- decide whether a proposal should be undertaken
- decide whether an existing project or program should be continued
- choose between different proposals (aimed at similar objectives)
- choose the appropriate scale and timing for a program or project.

However, cost–benefit analysis is often used when a program has matured or ended, as part of the evaluation of impact or outcomes. In this context the aim is either: to provide information on whether the outcomes achieved justify the resources used, relative to the alternatives; or to clarify the focus on different components of a project, to see how it may be improved, and which areas should be cut back or expanded. The most practical constraints in using CBA are the difficulties in assigning money values to the costs and benefits of a project or program (DOF 1991).

Software packages are available to assist in conducting CBA. One example is *Appraisal* – a package intended as a planning tool for scientists and program leaders to use in *ex-ante* analysis. The author (Appleyard 1996a) stresses that the process of attempting to produce a quantified measure of a project's benefits may be more valuable than the performance indicators generated. Evaluation of completed projects (*ex-post*) can determine whether expenditure on R&D was worthwhile, and may reveal benefits and costs that can be useful in future planning.

Key concepts of CBA

CBA compares a project scenario with an alternative scenario based on estimates of what would have happened in the absence of the project.

The following concepts are part of the cost–benefit analyst's toolkit:

- opportunity–cost – costs and benefits are priced at their value in their best alternative use, which may be above or below the actual cost of production
- willingness to pay – relatedly, costs and benefits are valued at what the marginal or last consumer in a competitive market is willing to pay for them
- the cost–benefit rule – a project is acceptable where, subject to budget restraints and certain other conditions, net social benefits valued according to the opportunity cost and willingness-to-pay principles are positive rather than negative.

CBA is often used in situations where the signals that the market normally provides are either absent or mistrusted: e.g. where no market exists, or inputs are underpriced relative to costs, or where outputs are overpriced. Various texts are available to provide guidance in these situations (DOF 1991). Costs and benefits in CBA should be valued at the specific time that they occur. In a retrospective analysis, costs and benefits are compounded forward to present value. This means that the cost–benefit rule in practice also needs to be consistent with the net present value rule:

- subject to budget constraints, a project should be undertaken if its net present value is positive; or
- subject to budget constraints, those projects should be undertaken with the highest net present values.

Sensitivity analysis is often used in association with CBA, because projects are exposed to many types of uncertainty. This usually involves testing the net present value (NPV) at 'best case', 'most likely' and 'worst case' assumptions, and can also be applied to the level of the discount rate (usually with a central value of 8%) (DOF 1991).

Cost–effectiveness analysis (CEA)

CEA differs from CBA in that benefits are expressed, not in money units, but in physical units (such as accidents reduced per unit of expenditure). Costs are expressed in money terms. It is often used in areas of health or education, where it is easier to specify benefits than to value them. It is useful in comparing alternative options or existing projects or programs. CEA does not provide an absolute measure of the benefit to the economy of the project.

Limitations and utility of CBA

The use of CBA is commonly criticised by agriculturalists on the grounds that unrealistic assumptions are often used in estimating costs and returns, and failure of the analyst to take account of important issues. However, agricultural extension professionals can often provide information that can greatly improve the estimates and assumptions used by economists in CBA. There is danger where the evaluation is focused on the bottom line in CBA, that the content of the analysis becomes obscured. On the other hand, conducting a CBA can be very enlightening for extension workers, and CBA can be a valuable tool in many forms of evaluation of extension projects. Most of the examples of CBA found in the literature are of analyses of research projects, or the combined effect of R&D projects. Case Study 3 provides an example of CBA conducted on an agricultural extension project 'Target 10' in dairying in Victoria.

Case Study 3
Cost–benefit analysis of the Target 10 dairy extension program

Evaluation	*Soft* ◄——————————●——————————► *Hard*				
Form	*Formative*			*Summative* ●	
Owen's form	*F5*	*F4*	*F3*	*F2*	*F1* ●
Data collection	*Qualitative*		*Both*		*Quantitative* ●
Program planning level	*Regional*		*State* ●		*National*

The program
Target 10 has the aim of enhancing the viability of the dairy industry through programs that profitably increase the consumption of pasture per hectare. Information from R&D on pasture utilisation is extended to farmers through courses, discussion groups, newsletters, target graphing, re-designed feed planning, comparative analysis, field days, focus farms and demonstrations and other media.

Goal of the evaluation
To estimate the economic benefits and to conduct a cost–benefit analysis of the expected results.

The evaluation and assumptions
This was a preliminary evaluation based on benchmark studies prior to the project launch in 1993. It does not consider all benefits from project activities not directly associated with the project aim. The first step in the cost–benefit analysis was to estimate the average amount of pasture consumed/ha and the amount of milk produced had this project not been implemented in 1991. These and other assumption made are listed below:

level of pasture utilisation: *6.5 tonnes/ha (range 4-9 tonnes in Victoria)*
production from increased use: *7% increase in milk results from 10% increase in consumption*
milk production in Victoria: *511 514 litres/farm on 8469 farms (Q_0)*
levy and transport charges: *2.3 c/l (levy) and 2.7 c/l (transport)*
milk price received (P): *24.5 c/l – levies & transport charges = 19.56 c/l (P)*
average cost of production: *14 c/l before implementing Target 10 (AC_0)*
increase in producer costs *$2000/farm/yr required for implementation of Target 10*

The maximum potential benefits resulting from Target 10 were calculated using the incremental profit method (Morrison 1993), i.e., from $P = (P*Q_1 – AC_1*Q_1) – (P*Q_0 – AC_0*Q_0)$ where Q1 (Q_0 plus 7% increase in production) and AC_1 are the Quantity of milk and Average cost with Target 10, respectively. The estimate of maximum potential per farm was $4728 and for Victoria was $40 038 500.

Method (the cost–benefit analysis)
The total nominal cost of the Target 10 Project was $9 198 000 from various sources. The spreadsheet model Appraisal *(Appleyard 1996b) was used for the cost–benefit analysis. The benefit of the project is 'accelerated adoption'. Appraisal was run twice: first using the adoption profile had Target 10 not occurred, and second using the adoption profile with Target 10. The Net Present Value (NPV) was calculated as the difference between the NPVs of those two scenarios. To calculate the BCR (benefit cost ratio), the difference between the 2 present values of the benefits was divided by the present value of the project's costs.*

Target 10 was designed to accelerate the rate of adoption of favourable innovations related to pasture utilisation (years from information being available to uptake by farmers). In addition it was expected to increase the level of adoption (percent of farmers). It was assumed that without the project, maximum adoption of the relevant practices would be 32%. In 1997 40% of Victorian farmers attended Target 10 training courses across Victoria. Some further assumptions and the results of applying the Appraisal model are summarised in the table below:

Assumptions	Without Target 10	With Target 10
Year adoption begins	8	2
Maximum adoption	32	40
Years to obtain maximum adoption	15	7
Maximum potential benefit	40 038 519	40 038 519
Cost benefit analysis		
Present value benefits ($)	57 404 678[a]	120 084 711[b]
Present value of costs ($)		6 763 653[c]
Net present value ($) (b-c)-a		62 680 033
Benefit-cost ratio		10.27
IRR (%)		140.15

A sensitivity analysis was conducted for a number of possible adoption profiles. The year adoption begins and the maximum level of adoption without Target 10 are of particular interest because these contain the greatest amount of uncertainty. The year to reach maximum adoption was held constant and the results are shown in figure 7:

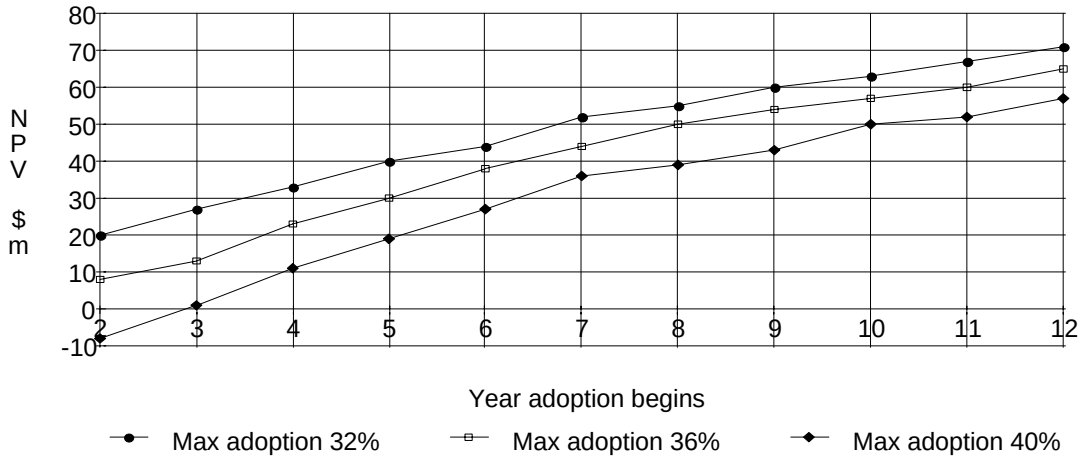


Figure 7 Graph of net present value at different adoption rates

The graph shows that the year adoption would have begun without Target 10 has a large effect on the project's outcome. The longer it would have taken for adoption to occur, the greater the benefit from extension work. The greater the acceleration due to Target 10, the greater the benefits. The graph also shows that the increase in maximum adoption resulting from Target 10 also affects the size of the benefits. A break-even analysis was conducted to work out the maximum costs that the producer could incur from implementing Target 10, before incurring a loss. Assuming a 7% increase in milk production from a 10% increase in pasture utilisation, this was found to be $5000 per farm.

Use of findings

Using the assumptions stated, the analysis indicates that extension has sped up adoption by 6 years, resulting in a net benefit of $62 million to the Victorian dairy industry. This preliminary analysis helped to validate project plans, and will be followed up by other evaluations by means of surveys of change in pasture use and farm profitability (Boomsma 1996).

Sources: Boomsma (1996); Appleyard (1996); Appleyard (1996a); Morrison (1993).

7.4. IMPACT EVALUATION FOR ILLUMINATION

Impact evaluation for illumination aims to explore both the intended and unintended outcomes of program interventions. Case Study 4 describes an evaluation of an overseas farming systems development project in India. Evaluation has a prominent presence in agricultural development projects in lower income countries. There is a strong focus on the use of logical frameworks, and an increasing tendency towards participatory process project design (Farrington and Nelson 1997). Often the word 'evaluation' is replaced by the term '**impact assessment'** (IA).

Internationally there is a current trend towards 'process' projects that have broadly-defined goals at the onset and are highly participatory. These projects pose particular difficulty for evaluation and require more of an exploratory approach in order to determine unexpected outcomes. Participatory process projects evolved as a backlash to the frequent failure of projects with pre-defined goals at the onset. These rigid programs were not flexible enough to meet the requirements of bottom-up planning. Case Study 4 is included in the Review as an example of one way of evaluating a 'process' project, where each site may have developed in individual ways and where summarising the results across the various sites would not be appropriate. The example uses village case studies to explore the impact of a participatory farming systems project conducted in Western India.

Case study evaluation is widely written about in program evaluation. In educational research and evaluation, case study is one of the main approaches taken. Internationally the value of conducting case studies is now recognised by development agencies such as the World Bank (Casley and Kumar 1987). The term 'case study' is used widely but loosely. Case studies focus on the individual but not necessarily on the singular. The individual is implicit in the notion of 'case', but the individual may be a group, a project, a village, an institution, a region and so on (Sechrest *et. al* 1996). Case study evaluation is the study of the particularity and complexity of a single case, coming to understanding its activity within important circumstances (Stake 1995). Case studies usually involve 'thick description'– that is description that is rich in context.

Sechrest et al. (1996) state that:

> "…a well-designed case study may provide justification for undertaking a larger scale study, and can provide critical information on 'how and what' questions. Because of the thick description and the many data points available, what a case study can do is reveal how the intervention works on the mediating variable(s) and therefore how the intervention brings change in the dependant variable. A case study may be the best design to answer questions about an emerging phenomena."

Case Study 4

Evaluation of a participatory process project in Western India

Evaluation	*Soft* ◄——●——► *Hard*				
Form	*Formative*			*Summative* ●	
Owen's form	*F5*	*F4*	*F3*	*F2*	*F1* ●
Data collection	*Qualitative*		*Both* ●		*Quantitative*
Stage of program	*Forming*	*Developing*		*Settled* ●	*Finished*
Program planning level	*Regional*		*State* ●		*National*

The program

The KRIBHCO Indo-British Farming Systems Project (KRIBP) as described elsewhere (Jones et al. 1996, Mosse et al. 1995) is a participatory farming systems development project situated in the Bhil tribal region of Western India. It was initiated in 1993 in one of the poorest regions in India. The project strategy, oriented towards improving the livelihoods of poor farming families, involves an extended process of participatory planning to generate location-specific plans for development of natural resources and social institution development. The project is based on the premise that sustainable development depends on enhancing local self-reliance through institutional and community development. The aim of the project model is to build participatory farming systems development approaches through the establishment of village level groups that target the very poor and take account of gender concerns.

The evaluative work relating to this paper was carried out between 1996 and 1997. However, it should be noted that in addition to this evaluation, the project conducted other types of evaluation. A monthly monitoring system is in place, in addition to a mid-term assessment of the impact of the project interventions, and assessment of the cost and benefits of participation.

Goal of the evaluation

The ultimate objective of the study was to determine project impact with respect to the objectives and aims laid out in the logical framework (see Section 12.2). In line with the project objectives special emphasis was given to the impact on the poor and women and on the participatory approach adopted by the project. Other aims of the study included:

- ***Documentation**: one of the aims of the project was to develop a replicable model for farming systems development. Because of this it was very important to document the project impact and experiences in order to feed this information into new projects and programs. This also helped to prevent loss of organisational learning and to provide suitable material for the orientation of new field workers.*
- ***Illumination**: donors and project staff wished to gain a fuller view of what was happening at village level, to understand how the participants perceived the combined impact of the various project interventions.*
- ***Accountability**: both project staff and donors were keen to be able to prove that this model was working, as prior to this there had been little documented evidence of strong project impact. The study was conducted towards the end of the first phase of funding of this project and was to be used to inform decision-makers as to whether this project should be continued.*

Form of evaluation

This evaluation focused more upon the impact than the process and has most in common with Owen's Form : impact evaluation.

Actors

The principal researcher had been attached to the project for a period of one year as part of a professional training scheme, and was under the supervision of the project management. However, the full evaluation team was made up of local field staff and international consultants.

Method of evaluation

Case studies of 8 project villages were conducted in order to understand the farming and livelihood systems and to get an overall view of the impact of the project initiatives. The case studies covered three states of Western India. In all studies, the project impact was assessed by considering local perceptions of project-induced changes, in the wider context of change occurring over the 5-year period in which the project had been operating.

During each case study household (HH) interviews, Participatory Rural Appraisal (PRA) exercises and semi-structured interviews (SSI) were planned to follow a logical order, to be flexible enough to enable the incorporation of learning as the process of information collection proceeded (see Figure 8). In line with the objective verifiable indicators (OVI), impact was considered with regard to who was benefiting: male or female; different socio-economic classes. Participatory farming systems models were used at an exploratory phase in this process, to ensure that the evaluation was based upon the values of the farmers themselves.

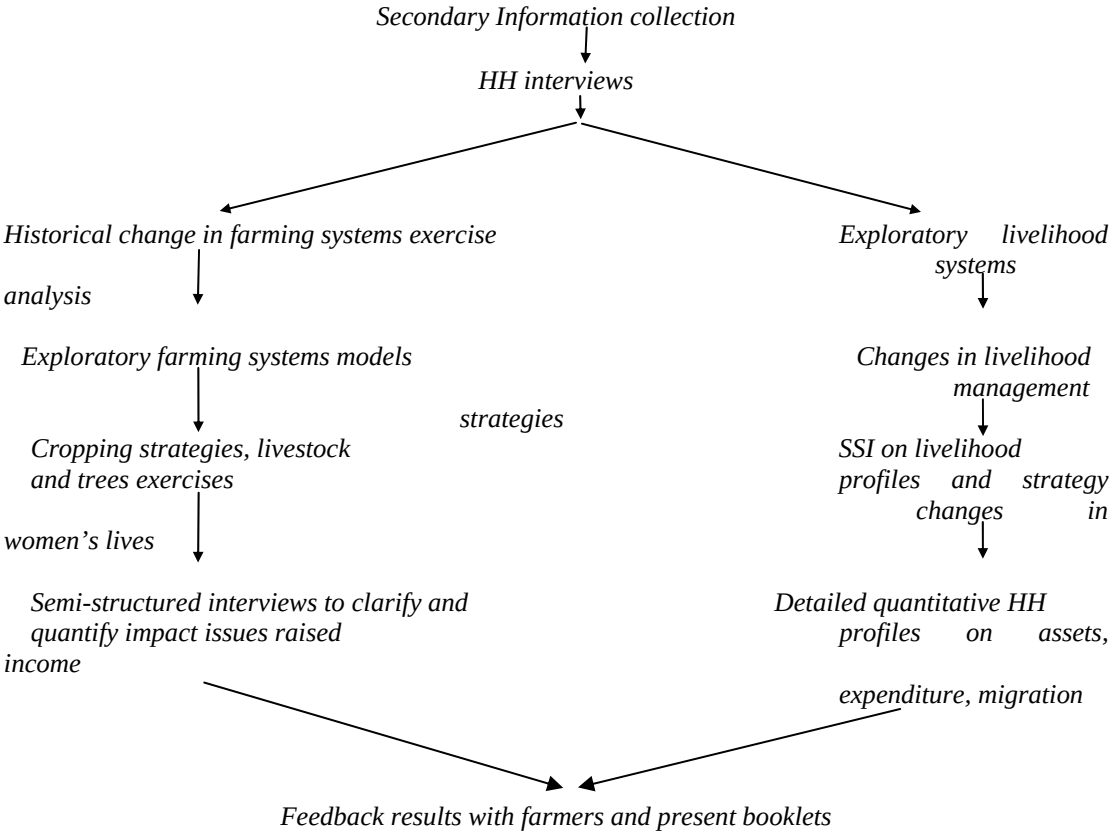


Figure 8 Information collection flow

Farming system models: *These models were used to learn how farmers perceived their farming system and to identify the sub-systems. It was hoped that this holistic view of the farming system would then create an ideal context in which to question the participants about any changes to this system brought about by the project. The participants were asked to represent the various enterprises of the farm using materials available nearby. In different villages the models were constructed using different materials, which included, clay, straw, seeds, twigs, dried cow dung and anything that was at hand. A key factor was that the facilitator was not involved in creating the model. When the diagram was completed the farmers were asked to draw arrows (or represent arrows with twigs) from one enterprise to another, to represent recycling and flows of materials and labour. Finally they were asked to represent the market*

and indicate flows of materials and labour to and from the market using arrows.
The second stage was to ask whether any changes had occurred in this farming system over the period of project involvement. Specific enterprises and relationships were pointed out and the participants were encouraged to talk about any changes in these flows and relationships. Every exercise was repeated and carried out with both male and female farmer groups. An effort was made to select informants from all socio-economic classes.

These models and other participatory exercises were used to explore the ways farmers view changes in their livelihoods and farming systems. Indicators such as the number of buffalo and the change in bride price came out of these exercises as local indicators of improved well-being. These were then used along with indicators developed by the project team to carry out semi-structured interviews. These findings were presented back to the farmers in booklets that included drawings, photographs and simple text, as the literacy rate was below 10 percent. During the feedback meeting the booklets were modified to include the farmers' comments and analysis of the results.

Use of findings
The participating farmers indicated that they were very pleased to receive feedback on the evaluation findings for a variety of reasons. Participants explained that they felt ownership over the evaluation and would keep the evaluation reports as village records for future use. Some of the women farmers perceived that being included in the process had elevated their status. While the findings did not feed directly into village work plans, it is hoped that they will be used to focus the next phase of village planning. The facilitators found this process to be an unobtrusive method to gain insight into a complex system; it also supported the participatory nature of the project. The case studies were documented and included many direct quotations from farmers as well as time-trend data for migration, herd size, tree cover, cropping intensity and change in crop variety.

The study revealed differences in impact among the villages selected in terms of project interventions that were chosen, and how they were implemented and received. The case study approach allowed in-depth analysis of specific situations and led to an understanding of unexpected impact. The approach to evaluation, especially the farming systems models, enabled farmers to identify what change factors were important to them. In this way, the methodology supported the overall participatory approach of the project. It allowed self-determination of which changes were important, and which changes should be measured. The evaluation also enabled project staff to further understand the values of the project participants, and this learning experience was fed back into the project's organisational learning system.
Source: Case study taken from Dart (1997)

8. FORM 2: MONITORING AND EVALUATION FOR PROGRAM MANAGEMENT

Monitoring and evaluation (M&E) for program management (based on Owen's Form 2) is designed to run throughout the life of a project and is generally set up during the developmental stage. This distinguishes it from other forms of evaluation, which are more discrete studies. The main focus of this form of evaluation is to provide feedback to the management information system. During the Review of Australian extension several organisations reported that they were initiating systematic monitoring systems, but we found little evidence of programs/projects that currently have systematic, formalised, established monitoring systems. This finding is confirmed by the conclusion of Woods *et. al.* (1993) that there is little evidence of formal M&E in Australia for the management of extension.

Despite this dearth of M&E within Australian extension, some Australian R&D workers have wide experience of projects in overseas countries, where such systems are established as an integral part of project design (McCormack, D. pers. Comm). Typically, the M&E system is derived from the indicators specified in a 'logical framework' (see Section 12.2) that is set up at the start of a project. A database is established to allow recording of M&E data on a time-sequence basis over the life of the project. The indicators are usually listed in four main categories, as shown in the example in Table 4.

Table 4 Example of matrix for recording M&E data for a rice development project over three years (shading indicates monitoring periods)

Indicators	Pre-project	Six-monthly time-series data					
	Dec 89	June 90	Dec 90	June 91	Dec 91	June 92	Dec 92
Financial							
project expenditure			▒		▒		▒
Physical							
no. of field days	▒		▒		▒		▒
no. of farmer contacts	▒		▒		▒		▒
sales of info. material	▒		▒		▒		▒
Benefits							
total rice production		▒		▒		▒	
average farm income		▒		▒		▒	
Environmental							
water use efficiency	▒						▒
soil phosphorus balance		▒		▒		▒	

From the data in the matrix, graphs are commonly produced that demonstrate progress in terms of project expenditure, extension activity, agricultural production and other indicators. These time-series data are intended to allow aid donors to keep track of expenditure and outcomes of the development effort. Gathering the data for an M&E system requires expenditure and effort, so simple and cheap and quantitative indicators are often employed. However, a sound M&E system requires carefully chosen indicators, which can include a range of more sophisticated measures, such as level of farmer satisfaction with project benefits.

In this Section several approaches to monitoring and evaluation are presented:

- planning-based approaches to monitoring using performance indicators (Section 8.1)
- evolutionary approaches to monitoring, that do not employ performance indicators (Section 8.2)
- quality assurance (Section 8.3)

M&E for program management is conventionally associated with input-output monitoring which would be represented by levels 1 to 3 of Bennett's hierarchy (Bennett 1977: see Table 5) and is explored further in Section 12.1.

Table 5 Bennett's hierarchy

Level	Bennett's Description	Type of evaluation
Level 7	Consequences for the target group	Evaluation & monitoring relating to social / economic impact
Level 6	Behavioural changes in the target group (Practice change, sometimes indicated by change in quality of products))	Evaluation & monitoring relating to behaviour change (also impact monitoring)
Level 5	Change in knowledge, attitude, skills, motivation and group norms	Evaluation & monitoring relating to intermediate indications of behaviour change
Level 4	The farmers' opinion about extension activities	Evaluation & monitoring relating to farmers' opinion
Level 3	Farmer participation in extension activities (participation monitoring)	Monitoring degree of participation
Level 2	Implementation of the program by extension agents (activities monitoring)	Activity monitoring
Level 1	Programming of the extension activities (inputs monitoring)	Input monitoring
Source: Modified from Bennett (1977)		

In Australia, many extension programs do record inputs and outputs for accountability purposes. But on-going records of farmer perceptions of extension activities, changes in KASA and behaviour change (levels 3-6) are more likely to be monitored informally; observations are written in extension workers diaries and internal files, or as verbal reports. A review of M&E of *Property Management Planning* programs in South-East Queensland showed that while a significant amount of information was collected for the program coordinators' own use, few data were reported formally. In some cases the data were collected for the team and reported

verbally, but not documented. The most frequent time period for reporting was six-monthly, and the type of information collected varied considerably from one coordinator to another. It was most common to monitor levels of participation in the meetings and farmer perception of workshop activities. Little information was routinely collected on changes resulting from the workshops (Stewart 1994).

The major recommendation of Purcell (1984) from a worldwide review of agricultural programs was that large expenditure of resources on extensive surveys to quantify extension impact and economic benefits is questionable for many countries, due to the extreme difficulty of linking cause and effect. He suggests that it is probably more appropriate to concentrate M&E on:

- measuring the change in adoption of ideas promoted by extension agents in specific areas of influence (KASA)
- understanding the reasons for significant non-adoption, and
- questioning the intended farmer clientele on how they perceive the extension services offered (by various agencies).

In other words it would be geared to improving effectiveness and efficiency of extension and directly asking intended clients about the impact it is having. It would be geared to ongoing evaluation rather than to quantitative ex-post evaluation, but would contribute to ex-post assessment.

8.1. PERFORMANCE INDICATORS USED IN M&E

Traditionally, formal monitoring in program management records performance (or management) indicators over a period of time. A performance indicator is a simple statistic recorded over time, to inform managers of the success of some aspect of program management. Examples of indicators in agricultural extension might be the number of farmers contacted by extension agents per year, or the number of members participating in discussion groups. Owen (1993) stresses the importance of ensuring that evaluators use a full range of data collection and analysis techniques. In terms of Bennett's hierarchy, a full range would include the above simple indicators, as well as some measures of change at higher levels, such as number of farmers adopting a new practice, or decrease in the number of farm families on social welfare. Fishpool (1993) listed five main categories of indicator:

- **Effectiveness indicators** – the extent to which a program is satisfying the purpose for which it was established. This type of indicator should provide answers to the question, "What sorts of things would you point to as evidence that the program was achieving the desired results?"

- **Social justice indicators** – the social impact of a program in terms of equity, equality, access and participation. This indicator should provide answers to questions like, "Does the program address all client target groups equitably?" "Is access (e.g. geographically, language, timing) to the program facilitated for all client target groups?" Or, "does this program increase the ability of clients to participate fully in the community or industry?"

- **Operational efficiency indicators** – the relationship between inputs and outputs. This includes both productivity and aspects of service delivery and should provide answers the questions like, " how much does it cost to service our clients and how does this compare to previous performance or to the costs of similar operations elsewhere?"

- **Outcome efficiency indicators** – the relationship between outputs and outcomes. This indicator should provide answers to questions like " how can we examine the costs against the benefits of this program?" or "how do we know whether this program achieves its outcomes efficiently or are there better ways?"

- **Standards of service indicators** - the quality of the service to clients. This indicator should provide answers to questions like "How do we know if our clients are satisfied with the way they are treated when using the program? How do our current standards of service compare with the past levels of services?" Or, "is the standard of service provided appropriate to client needs?

The danger is that simplistic performance indicators can miss out on the bigger picture. In addition to this, reliance on quantitative indicators alone can 'lead' the program: they are not neutral and can be counter-productive and even run against the greater aims of the program. There is now a significant debate about the past usefulness of performance indicators in government agency and corporate planning. Owen (1993) cites Winston (1991):

> "It appears that there is no evidence in the literature or in the practical experience of governments that performance indicators have ever been successfully applied by governments to the evaluation of human service programs, in the context of program budgeting. There is evidence to the contrary, that exemplifies failures to get performance indicators to contribute as intended to management and budget cycle decision making" (Winston, 1991:604).

The following case studies represent M&E studies of various types. It is significant that although many good examples of M&E exist in health and education, no documented evidence was found of comprehensive M&E in agricultural extension. Case Study 5 illustrates a monitoring system, using indicators that correspond with levels 1 to 4 of Bennett's hierarchy. However, the reader is advised that in this case study monitoring of crop pests is also intervention oriented and can cause some confusion. No other suitable case study was found of monitoring extension programs in Australia.

Case Study 5
IPM program for potato pests: an example of monitoring & evaluation

Evaluation	*Soft* ◄——●——► *Hard*				
Form	*Formative* ●			*Summative*	
Owen's form	*F5*	*F4*	*F3*	*F2* ●	*F1*
Data collection	*Qualitative*		*Both* ●		*Quantitative*
Stage of program	*Forming*	*Developing*		*Settled* ●	*Finished*
Program planning level	*Regional* ●		*State*		*National*

The Program

Until 1987 the Australian potato industry relied heavily on chemical insecticides; the regular use of organophosphates and synthetic pyrethinoids remained common practice in many areas (Horne 1990). The Horticulture Research and Development Corporation (HRDC) is keen to promote Integrated Pest Management (IPM) amongst potato growers. The aim in this project was to extend information on IPM to potato growers, crop advisers and others in the industry by demonstrating IPM in action. Evaluation of the extension program was built into the project and comprised monitoring both the uptake of IPM by farmers and the monitoring of pests in farmers' crops. However, this crop monitoring was also an essential part of extension and of the IPM concept itself (Hall 1994).

Regular crop monitoring is an important component of IPM because it allows growers to make timely decisions with precision on action to be taken against pests. This aspect of monitoring was done by trained and trusted crop advisers, and the data were used by farmers as well as to provide information (with other indicators) for on-going evaluation of management of the extension project.

In the areas selected for the project, support from growers who owned the monitored crops was important in showing that IPM was practical and could be used immediately to provide good results. Information based on these demonstration crops and their pest populations was presented at grower meetings, field days, industry meetings, conferences, workshops and in written articles.

In Victoria, a small number of growers (4-10) in each of 3 key districts cooperated with the field monitoring and demonstrations. Parallel projects were conducted in NSW and Western Australia, Queensland and South Australia. In May 1997, a protocol for monitoring potato crops in Australia was developed by bringing together people in the project who were involved in advising growers. This provided a procedure that can be used as the basis for monitoring in any potato crop in Australia.

Goal of the evaluation

The main aim of M&E was to provide continuous feedback to the project managers to allow constant improvements to be made towards the project goal of promoting IPM amongst growers.

Form of evaluation

This case study is an example of 'monitoring and evaluation' for management. However, the methods of data collection were varied and rather informal – because of the difficulty of establishing an extensive infrastructure for regular recording of adoption and other performance indicators. A range of techniques and types of information were utilised in the first year. In the longer term it may be feasible to establish more objective and quantifiable indicators. The researchers feel that excessive time may be wasted in establishing too extensive and rigid an M&E system, and that informal means of M&E may be more time- and cost-effective.

Actors
The evaluators were members of and collaborators with the project team

Methods of evaluation
A number of measures were proposed as indicators of the progress of the project towards achieving adoption of IPM by potato growers. These included 'knowledge of principles of IPM', 'number of insecticide sprays per crop (for particular pests)', 'knowledge of beneficial insects in potato crops' and 'percentage of farms adopting IPM practices'.

In 1987 there was virtually no knowledge or use of IPM in the Australian potato industry. Prior to this project commencing in 1996, mail surveys were distributed to (2356) growers in all major potato growing areas across Australia to determine the level of awareness or use of IPM at that time. The 680 growers who responded indicated wide variation in awareness of IPM - from 35% in some districts to around 60% in others That survey served as a baseline for the M&E of data from various districts in this extension project. Data on these indicators are collected from periodic surveys and advisers in key districts, from processors and other sources. These data are examined together with other data from each district, on 'cost of insecticides used per crop', 'yield of potatoes per hectare', 'quality and price of potatoes at marketing' and other variables useful in assessing the performance of the project. Time sequence surveys have been conducted in some key potato growing areas (e.g. Ballarat) and these indicate progress towards adoption of IPM. Anecdotal information from growers on their experience in the use of IPM and of the extension program is also recorded in all districts.

Some results
The M&E system established appears to be providing information valuable in adjusting project management towards achieving project aims of extending the use of IPM. Insecticide use is declining in many areas and crops have produced yields at least as high as usual (Paul Horne, 1997, pers. comm). Support from growers as a result of these demonstrations was very high and some selected responses of growers were given in an article in Good Fruit and Vegetable *(June 1997) and in O'Sullivan (in Eyes on Potatoes. In general, growers found that IPM resulted in significant savings by reducing the number of insecticide sprays, without loss of production or quality.*

.
It is has become clear that IPM strategy must be adaptable to suit local conditions and production methods. For example, fewer insecticides are being used on monitored crops grown under centre pivot irrigation near Swan Hill, despite higher pest numbers than in other potato growing districts.

The project has led on to the development of specific IPM solutions to particular problems, such as the instigation of commercial breeding of beneficial Orgilus *wasps for release in selected areas. Another important outcome for management has been the Monitoring Proforma developed for the project, which is now used by many advisers and consultants to serve growers in many districts*

Use of findings
The results are utilised continuously in improving the management of the project and to guide others in establishing IPM programs. Using the results to date, support is being sought to expand the project to other districts.

Source: Paul Horne (1997) IPM Technologies P/L, Victoria

8.2. EVOLUTIONARY APPROACH TO PROJECT MONITORING

Current approaches to monitoring extension programs overseas are heavily influenced by a planning ethos that places substantial emphasis on rationality, prediction and control (Davis 1994). The evolutionary approach to monitoring described in Case Studies 7 and 8, in many respects, is the opposite. Davis (1994) outlines seven main differences between these approaches:

- **Objective indicators vs subjective iterative process:** A central feature of planning-based methods of monitoring is the use of 'indicators'. Within planning-based approaches it is believed that differences in the subjective perspectives of events, and the underlying value concerns of different observers, need to be controlled or ignored. Under the evolutionary approach, agreement on meaning of events is an outcome at the end of a process (a month's cycle or more), never final in its form, and subject to revision in the light of new experience. The identification of differences in interpretation is a central to the whole process; they are to be brought to the surface and explored, not ruled out.

- **Inclusion vs selection:** Planning-based monitoring systems are mostly quantitative in their content. Quantitative analysis is based on the ability to enumerate things or events. At the very basic level of counting, quantification is also about the homogenisation of experience. Within the daily experience of organisation those events which are countable are summarised by a process of inclusion. Within the evolutionary approach experience is summarised by selection rather than by inclusion, it focuses on the exceptional rather than the commonality, and it seeks to differentiate rather than homogenise. It is about defining the meaningful edges of experience rather than identifying a central tendency.

- **Predicting outcomes vs open-ended outcomes:** Under the planning-based approach to monitoring, events of concern are identified before their occurrence, rather than afterwards. In conventional systems 'indicators' are established at the beginning of a project, and data in the form of statistics is gathered repeatedly throughout the life of the project. The process is strongly deductive in orientation: starting with a conception of the desired outcomes and working down from there. The opposite is an inductive approach, where indicative events are abstracted out of recent experience, and this process is renewed with each new reporting period of the monitoring system. Instead of being predictable it is open-ended.

- **Who defines the indicators?:** In most monitoring systems events of concern are defined by people distant from those events that are to be monitored. Typically the identification of indicators is carried out by senior staff in organisations. Reformist approaches have consisted of taking the indicator identification process down the hierarchy, in some cases, to the beneficiaries themselves whose views are sought, through the use of PRA methods. In evolutionary monitoring those closest to the experience being monitored (e.g. the field staff) are given the right to pose to those above them a range of competing interpretations of those events. The role of those in power over them then becomes to respond, on a selective rather than inclusive basis, to the menu of options provided to them a new each month.

- **Where are the data analysed?:** Normally the analysis of events documented by an organisation monitoring system is carried out on a centralised basis, at senior levels of the organisation. Typically, field level workers do not analyse the data they collect; rather they simply forward information up their hierarchies for others to analyse. In evolutionary monitoring, information is not stored or processed on a centralised basis, but is distributed throughout the organisation, and processed locally. Staff not only collect information about events but they make their own evaluation of that information, according to their own local perspective.

- **Statistics vs thick description:** Normally when conventional monitoring data is analysed it is in a form and location that strips it of context. Typically, few text comments accompany statistics sent up from field workers. The alternative makes use of what Geertz (1973) has called 'thick, description', closely-textured accounts of events, placed in their local context, and where the role and subjectivity of the observer is visible. In the world of ordinary people these often take the form of stories or anecdotes. Within the evolutionary approach to monitoring outlined here these 'stories' are accompanied by their readers' interpretations.

- **Static vs dynamic:** Most monitoring systems are largely static structures. Indicators remain essentially the same for each reporting period, and the same questions are asked again and again. The focus remains the same. With the evolutionary approach the contents of the monitoring system are potentially far more dynamic and adaptive. Events reported reflect both a changing world and changing sets of perceptions within the organisation about what is important within their world. Where quantitative data are sought on the incidence of an event found to be significant by a number of levels of the organisation, it can be done on a one-off basis; there is no intrinsic need to repeat the same inquiry each reporting period thereafter

(**Source: paraphrased from Davis 1994**).

Case Study 6 is an example of an innovative approach to monitoring without indicators from the field of agricultural extension in Australia: Case Study 7 is included to illustrate an evolutionary approach to project monitoring in an overseas development project.

Case Study 6

FARMSCAPE (Farmer-Adviser-Researcher Monitoring, Simulation, Communication and Performance Evaluation) for improved management of dryland cropping

Evaluation	*Soft* ◄——●——► *Hard*				
Form	*Formative* ●			*Summative* ●	
Owen's form	*F5*	*F4*	*F3*	*F2* ●	*F1*
Data collection	*Qualitative* ●		*Both*		*Quantitative*
Stage of program	*Forming*	*Developing*		*Settled* ●	*Finished*
Program planning level	*Regional* ●		*State*		*National*

The project

FARMSCAPE is currently funded for 3 years and commenced in 1995. It is run by APSRU (Australian Production Systems Research Unit), a group made up of officers from the CSIRO and Queensland DPI. APSRU is directly interested in the geographical area from Northern New South Wales to Central Queensland. APSRU's core technology is based on a computer program, Agricultural Production Simulator. *APSIM uses meteorological, crop and soil data to simulate the consequences of management on system performance. FARMSCAPE's decision support features a high degree of contextualisation and provision of insights rather than prescriptions. The notion of decision support has evolved from a sole emphasis on developing software as a ready-made product for use by farmers, to developing new relationships and learning processes.*

The FARMSCAPE project (that this evaluation concerns) was developed from a pilot project established in 1992 as part of a process to link APSRU with 3 farmer groups operating in the Dalby region. The aim was to demonstrate tools that were available to improve sound decision making, as well as providing feedback on those decision-making areas most likely to benefit from the products of research and modelling. The FARMSCAPE project had been operational for 2 years when this evaluation commenced.

Goal of the evaluation

The overall goal of the evaluation was to determine the degree to which various aspects of the FARMSCAPE RD&E approach was contributing to how managers think about management of their production activities in the face of extra-ordinary weather uncertainty, changes in their practice, how this was happening and in what ways it could be enhanced. More specifically it aimed to:

- *identify opportunities for contributing to new insights and better planning and decision-making*
- *log evidence of the degree to which a project activity contributed to new insights and better planning and decision-making.*

Form of evaluation

This evaluation is best described as Owen's Form 4: monitoring and evaluation for management. It is an on-going evaluation and is aimed at improving the delivery of the program as well as contributing to making summative judgements at a later stage. This evaluation process could also be described as Form 2, evaluation in program management.

Actors

Both internal and external evaluators carry out the evaluation. The audience appears to be mainly internal to the program, including the farmer clients.

Method of Evaluation
The central evaluation tool was the use of 'logs'. In the context of this study, logs are iterative records/interviews of events, reactions, attitudes and action captured in a given framework, through a combination of:

- *individual written notes/reports by key persons at selected intervals, or as an issue arises*
- *structured debriefing sessions by selected groups themselves or facilitated by an outsider*
- *semi-structured interviews with key persons.*

Currently, 30 persons are 'logged' by telephone interview every 3 months and are interviewed face-to-face by an external evaluator every 6 months. The human logs include farmers, consultants, DPI extension staff, researchers and modellers. This method of evaluation is still being developed. The intention was to trial various approaches and to select the most effective ones to use. The logs are to be used with each of the main stakeholder groups in the project.

Presentation/utilisation of findings
It is intended that the information gained in this process will be fed back to project members to assist in improving future actions.

Source: From information supplied by R. McCown, CSIRO (Australian Production Research Unit, Toowoomba) and J. Coutts (Rural Extension Centre, University of Queensland, Gatton)

Case Study 7

An evolutionary approach to facilitating organisational learning: An experiment by the Christian Commission for Development in Bangladesh

Evaluation	*Soft* ◄●					►*Hard*
Form	*Formative* ●			*Summative*		
Owen's form	*F5*	*F4*	*F3*	*F2* ●	*F1*	
Data collection	*Qualitative* ●		*Both*		*Quantitative*	
Stage of program	*Forming*	*Developing*		*Settled* ●	*Finished*	
Program planning level	*Regional*		*State*	*National* ●		

The program

The Christian Commission for Development in Bangladesh (CCDB) is a medium-sized Bangladeshi non-government organisation (NGO) with almost 550 staff. Its annual budget is approximately US$4 million. The main program CCDB runs is the Peoples' Participatory Rural Development Program (PPRDP), which involves more than 46000 people in 785 villages in 10 districts. Approximately 80% of the direct beneficiaries are women. Development assistance is made available to participants in 3 forms: group-based savings and credit facilities used to meet the needs of individual households, grant assistance given to the same groups on a pro-rata basis and intended for community level developments, and skills training, mainly for livelihood purposes. The large scale and open-ended nature of these activities poses a major problem for the design of any system intended to monitor process and outcome.

Goal of the evaluation

In 1994 an experiment in participatory monitoring was conducted with PPRDP program that involved the deliberate abandonment of the use of 'indicators', a central concept in orthodox approaches to monitoring. Implementation took place in 4 PPRDP project areas in Rajshahi zone of western Bangladesh, where 140 project staff are working with approximately 16,500 people grouped into 503 associations.

Form of evaluation

In terms of the meta-model of Owen (1993), this form of evaluation best fits under Form 2; monitoring and evaluation in program management.

Actors

The experiment was the outcome of a voluntary collaboration between an independent researcher (external) and the program (CCDB). CCDB adopted the approach because, according to the Director, it appeared to fit their needs. The independent researcher also had a need to generate information of value to a PhD thesis. The researcher was not accountable to the donors.

Method of evaluation (an outline of the process, as implemented)

The selection of domains of change to be monitored: *A brief process whereby senior staff identified 3e broad areas or types of changes they thought CCDB needed to monitor at the project level:*

- '*changes in people's lives*'
- '*changes in people's participation*'
- '*changes in the sustainability of people's institutions and their activities.*'
- '*any other type of change*'.'

None of these types of change were precisely defined. Their boundaries were deliberately left 'fuzzy'. Initially it would be up to the field level staff to interpret what they felt was a change belonging to any one of these categories. Since the first trial of the method in March 1994 changes have been reported for each month in each of the 4 project areas.

The participants: *There were 4 groups of participants in the monitoring system: (a) group members in the project area, (b) the local project staff, (c) the senior staff at a central office (d) CCDB's donors. The structure of their participation determined how the information from the monitoring system was analysed.*

Phrasing the question: *The basis of the monitoring system was a simple question in the following form: "during the last month, in your opinion, what do you think was the most significant change that took place in the lives of people participating in the project?" The respondent was then asked to give an answer in 2 parts. The first part was descriptive: what happened, who was involved, where it happened, when it happened. In the second part the respondent was asked to explain why they thought the change was the most significant out of all the changes that took place in that month. In particular, what difference had it made, or will it make in the future? The monitoring system was not aiming to report the average state of the project, but on the most significant events.*

The structure of participation: *In March 1994 a workshop was held with the senior staff of the 4 local project offices to plan the implementation of the monitoring system. Each Project Office was told that at the end of each month thereafter they would be expected to report to the head office one significant change in each of the 4 domains of change. Each Project Office was then asked to draw up a plan for how their field staff would, each month, identify a range of potentially important changes and how these would then be analysed in order to identify the most important. In practice, an average of 15 changes were documented by the field staff at each Project Office, each month, out of which 4 were then selected by more senior staff in the same Project Office as the most significant and then sent on to Dhaka.*

The process was repeated at the Dhaka head office. The 4 sets of 4 changes (one set of 4 from each Project Office) were brought to the head office each month. The task of the head office staff was to select the 4 changes from the 16 which they thought were the most significant of all. In practice between 4 and 8 senior staff attended each of the monthly meetings and decided that each participant would rate each story out of 10, and the ratings would then be aggregated to produce the group response. The rating process was preceded by an active group discussion of each account of change. The single requirement was that they must document and explain their choice and communicate it back to the staff in the 4 Project Offices. In practice, the Dhaka office meeting and discussion took about 3 hours of staff time per month.

The third level in this process was analysis was by the donors. By the end of September the head office had selected 24 accounts of significant changes (4 domains x 6 months). Those changes were collated in the form of 4 chapters in a report. It was proposed that donors should read each chapter and select the one change in each chapter that they thought was the most significant according to their own values and concerns. As with other participants, they should document the reasons for their choices

Feedback: *After each month's changes were evaluated by the head office staff, their judgement of the most significant changes, and the reasons behind those judgements, were fed back to the Project Offices concerned. Similarly, the results of the sub-group discussions at the RTM were also fed back via a plenary session. The purpose of regular feedback was so that those identifying the changes in the first instance could take into account the views of CCDB senior staff when in the process of evaluating changes. It was intended that if feedback was provided as planned the monitoring system should take the form of a slow but extensive dialogue up and down the CCDB hierarchy each month. In more evolutionary terms it can be seen as a process of co-evolution of interpretative frameworks within an organisational ecology.*

Verification: *Those changes that were identified as the most significant of all were precisely those stories where the most effort needed to be invested in verifying the factual details of the event. Verification visits to the sites of the described events can perform a policing function, ensuring that field staff are kept honest in their report-writing. They also provide an opportunity to gather more detailed information about the event which was seen as specially significant, and if some time after the event, a chance to see what has happened since the event was first documented.*

Quantification: *This can take place at 2 stages. Firstly, when an account of change was being described it was quite possible to include quantitative information as well as qualitative information. Secondly, it was possible to quantify the extent to which changes identified as the most significant in one location or zone have taken place in other locations, within a specific period. In the case of one significant change identified (concerning a women's purchase of land in her own name) all 10 Project Offices were asked to supply information on the numbers of identical incidents that they were aware of having taken place in their project area in the past year. However, no need was seen to repeat this particular question every month thereafter, as in traditional monitoring systems.*

Use of findings

In summary, the monitoring system has survived and is itself undergoing evolution, both in the specifics of its procedures and in its perceived purpose. Although initially planned to operate for the 6 months until the RTM meeting the monitoring system described above was continued afterwards, on the instructions of the Director. In January 1995 CCDB decided on its own initiative to extend the system to include 3 more of the 10 PPRDP areas. Rather than the contents becoming increasingly similar over time, and the system reaching a form of steady state, new changes have continued to be reported. The most notable of these concerned association involvement in an intra-family conflict over contraception use, reported in December 1995. Changes also occurred in the success ratings of the different project centres; while the Director had previously identified one Project Office as the most successful, on the basis of its good credit repayment record, the same Project Office was the least successful in terms of its ability to generate, through the PMS, a large number of highly-rated accounts of significant change. Staff have made extensive use of stories of change in CCDB publications, videos and educational materials. In addition, Project Office staff took visiting donor representatives to associations featuring the reported significant changes immediately prior to the RTM in November.

Source: Rick Davies, Centre for Development Studies, Swansea SA2 8PP, Wales, URL: http://www.swan.ac.uk/cds/rd1.htm

8.3. QUALITY ASSURANCE

At the beginning of the Review, its scope was defined as being concerned with 'programs' rather than product, policy or people. In extension, however, 'product' evaluation and 'program' evaluation sometimes overlap. Quality assurance may be thought of as 'product' M&E, but within the field of agricultural extension, performance indicators are often associated with product quality and thus in this context quality assurance is inseparable from program evaluation. Owen (1993) includes quality assurance in the Form 2, as it is primarily for program management. Quality assurance is motivated by the need of governments, industries and competitive companies to be seen to be delivering high quality services with minimum expenditure of funds. It is becoming an increasingly important concept in all agricultural industries in Australia for products aimed at both export and internal markets.

Case Study 8 is an example from the poultry industry, in which private and government extension agents became involved in design and implementation (and training) for a QA system.

Case Study 8
Quality assurance for Farm Pride Food, Victoria

Evaluation	*Soft* ◄				●				►	*Hard*
Form	*Formative*		●				*Summative*			
Owen's form	*F5*	*F4*		*F3*			*F2*	●	*F1*	
Data collection	*Qualitative*			*Both*	●			*Quantitative*		
Stage of program	*Forming*		*Developing*	●		*Settled*			*Finished*	
Program planning level	*Regional*			*State*	●			*National*		

The program

Andrew Almond (poultry training and extension specialist) was commissioned to coordinate the design and implementation of a Quality Assurance (QA) system for a major egg producer cooperative – Farm Pride Food. This activity is part of his extension activity, which aims to improve the general performance and profitability of the poultry industry. Farm Pride Food markets 90% of all eggs in Victoria and has 110 members ranging in size from 200 to 240 000 layers. The cooperative's major customers (Coles, Woolworths, McDonalds) have instituted their own QA programs and consequently are insisting that all their suppliers implement QA too. Simultaneously the growing consumer concern about food safety issues has led Farm Pride Foods to put together a Food Safety Plan. These 2 pressures led to a decision by Farm Pride to design and implement their own QA in 1996/7. Industry/DPIE funding was procured to support the development of a QA system.

Goal of the evaluation

The aim of Farm Pride was to secure the position of preferred supplier status with their major customers. This intention was combined with the goal of reducing food safety risks to a minimum. In the longer term, the system is intended to lead to ISO 9002 status for the industry and some producers, which is required for export marketing of egg products. For Farm Pride, QA was implemented in an attempt to ensure higher quality management and products from their suppliers and to reduce safety risk. For the egg producers the system enables complying members to gain accreditation and all the marketing advantages of QA status.

Form of evaluation

QA is considered to be a form of M&E for management because it involves setting aims and standards and monitoring compliance.

Actors

Both external and internal auditors are used now that the system is established. The program was commissioned by Farm Pride and developed jointly by external specialists with their own staff. The training program involved a number of external specialists, and training manuals and other documentation were developed jointly with participation of producers. The audience is both the Farm Pride cooperative and the individual producers, as well as their customers.

Method of evaluation

A flow chart of steps and activities (Figure 9) shows the process adopted in establishing this QA system over a period of 18 months from February 1996 until now. The QA system comprises 3 elements:

- *establishing a Code of Practice (general rules for production of first quality eggs)*
- *good Management Practices (GMP); a guide for farm managers to produce good quality eggs. includes grades and standards, e.g. for packaging and transport*
- *HACCP-based food safety plan (Hazard Analysis at Critical Control Points – a world standard procedure for analysing food supply standards).*

The main steps in the design and implementation of the QA system are shown in the flow chart below:

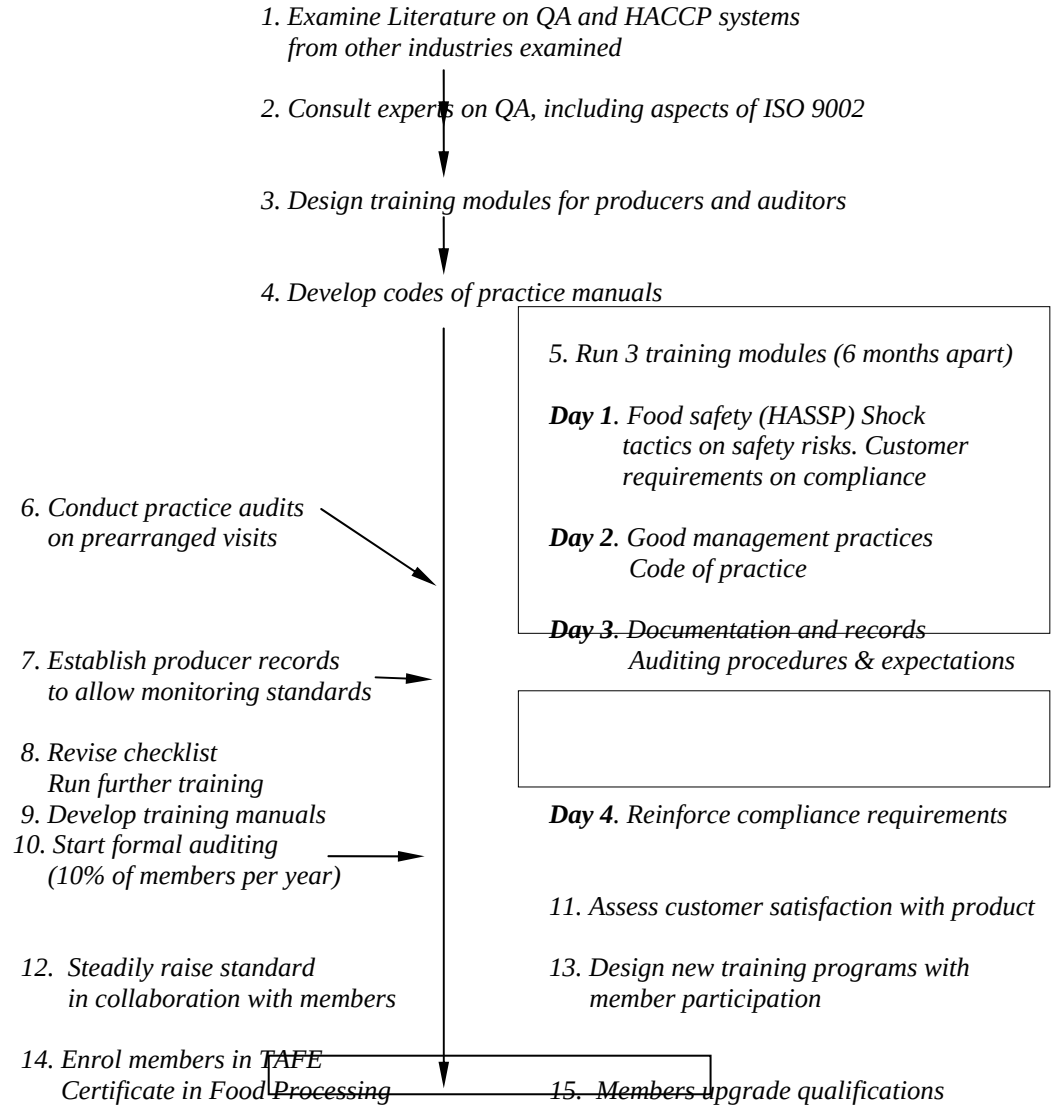


Figure 9 Flow chart of activities in establishment of a QA program for Farm Pride Foods

Use of findings

The QA system is well accepted by members and has been 'bought' by Farm Pride Foods in other states. The compliance rates by producers has increased markedly over the year since training began.

Source: Drafted from information received from Andrew Almond, Longerenong College (University of Melbourne) and Nagi Mina, Farm Pride Foods, Keysborough

9. FORM 3: PROCESS EVALUATION

The principal focus in process evaluation is on the improvement of a program and understanding how it is being implemented. This form is generally carried out at before program completion and is distinct from monitoring in that it is a discrete activity. Examples of models include implementation studies, action research, responsive evaluation, and topical RRAs (rapid rural appraisal – see Section 11.3) or PRA (participatory rural appraisal – see Section 11.4). The Review of evaluation in agricultural extension has found several approaches to process evaluation, including:

- process evaluation that is carried out to understand how a program is being implemented (illumination evaluation)
- process evaluation that is particularly responsive to the needs of the program clients (farmers) and aimed at improving the program while it is still developing (responsive evaluation)
- process evaluation that is carried out to provide guidance in the refinement of a program. This approach is often carried out when reviewing pilot programs, and has a strong formative element.

9.1. PROCESS EVALUATION CARRIED OUT TO UNDERSTAND PROGRAM DELIVERY

During the 1960s and 1970s the importance of examining implementation in formal evaluation studies emerged. When programs failed to achieve desired outcomes, practitioners started to examine how the programs had been implemented. It was found that the type of implementation varied greatly from site to site and that the process of implementation was an important factor in determining whether desired outcomes would be reached. Studies of this kind are referred to as **process-outcome studies** and are generally summative in nature (Owen 1993). However, in some cases the end point of the process-implementation study is to simply assess the extent of implementation. Case Study 9 provides an example of a process evaluation that was conducted to understand the degree of implementation of the program and to investigate stakeholders' perspectives.

Case Study 10 is also an example of evaluation carried out to understand the delivery of the program but is more directed towards illumination. They also differ with respect to the level of organisation that is considered: unlike Case Study 9, Case Study 10 is aimed at evaluation of the progress of individual groups rather of a national program.

Case Study 9

National Whole Catchment Management: review and analysis of processes

Evaluation	*Soft* ◄	●	———	►	*Hard*
Form	*Formative*		*Summative*		
Owen's form	*F5*	*F4* ●	*F3* ●	*F2*	*F1*
Data collection	*Qualitative* ●	*Both*		*Quantitative*	
Stage of program	*Forming*	*Developing* ●	*Settled*	*Finished*	
Program planning level	*Regional* ●	*State* ●		*National* ●	

The program

The Land and Water Resources Research and Development Corporation (LWRRDC) in partnership with Murray-Darling Basin Commission and the National Landcare Program (NLP) established a national program on dryland salinity research, development and extension. The program seeks to develop a national response to dryland salinity that optimises the use and sustainability of natural resources. The principal aim of the program is to foster the development and implementation of catchment management plans which are closely linked to state salinity and/or whole catchment planning strategies. The program intends to support activities related to the priorities and recommendations outlined in catchment management plans that have been developed at the community level with state agency assistance. For purposes of this evaluation, one focus catchment was selected in each of the 5 mainland States. The outcomes and experiences gained in these catchments were to be extrapolated to other regions.

Goal of the evaluation

The goal of the evaluation was to examine WCM (Whole Catchment Management) at a national level, with a view to ensuring that the National Dryland Salinity Program is adequately integrated with sustainable environmental management at a catchment level. To do this a generic procedural model for ensuring that participatory and coordinated WCM occurs was developed through this evaluation.

Form of evaluation

This evaluation best fits Owen's Form 3, process evaluation. It is largely directed towards a study of program implementation and is aimed at increasing understanding and improving decision-making from the catchment to the national levels. It is formative in nature and is directed towards describing the program and developing a framework for further evaluation.

Actors

An external evaluator was commissioned by LWRRDC to carry out the evaluation. While the client was LWRRDC (internal) the audience was much broader – from landholders and Landcare groups to natural resources agencies.

Method of evaluation

Throughout the document the various stakeholder views were represented separately under the consideration that all views are valid (a multiperspective approach). Case studies of the catchment areas operating under the WCM program were carried out. An analysis of state structures and progress towards WCM on specific catchments was undertaken. Selective reviews of the available catchment management literature were conducted, followed by interviews with a wide range of stakeholders and site visits to a range of catchments on a national level. Interviews were conducted by telephone and face-to-face during site visits. Structured questionnaire were not used and data analysis was qualitative. Concerns and issues raised in the interviews were paraphrased and grouped under appropriate sections.

Use of findings
The state-by-state analysis was used to identify stakeholders in WCM and derive a series of principles for WCM, which were suggested as best practice nationally. The principles were then used to derive a recommended generic Procedural Framework, PRIME, in which the planning, research, implementation, monitoring and evaluation activities essential to the achievement of WCM are described. In the report the findings of the evaluation were not presented as a series of specific stand alone recommendations, but as a 'package', with the procedural model (PRIME) being appreciated as a logical derivative of the WCM principles identified nationally.

Source: From materials supplied by J. Butterworth and G. Syme, CSIRO Land and Water, Perth

Case Study 10

Corangamite Salinity Program and global assessment scaling

Evaluation	*Soft* ◄	●	———	———	———	► *Hard*
Form	*Formative*	●		*Summative*		
Owen's form	*F5*	*F4*	*F3*	●	*F2*	*F1*
Data collection	*Qualitative*	●	*Both*		*Quantitative*	
Stage of program	*Forming*		*Developing*	●	*Settled*	*Finished*
Program planning level	*Regional*	●	*State*		*National*	

The program

The Victorian DNRE decided to have a rigorous look at the effectiveness of one of its programs directed against the problem of dryland salinity. This program had been running for 5 years, and was considered to be at a settled stage. The specific project to be evaluated was the Corangamite salinity program. This case study describes an evaluation that was carried out to complement the existing road-side surveys (monitoring physical environment by observation) with some monitoring of institutional building and group forming.

Goal of the evaluation

The aim was to monitor the progress of the community groups involved in the program and to determine whether group performance was related to the amount of salinity control activities undertaken. This was to be conducted with an overall goal of understanding how the existing project was functioning and to improve the delivery of the project. A further goal was to test some new methods to augment the department's basket of choices of evaluation tools and approaches.

Form of evaluation

This evaluation best fits Owen's Form 3; process evaluation. It was conducted with view to further understanding the implementation of the project rather than solely concentrating on the physical outputs /impact.

Actors

The evaluator/facilitator was internal to the department and had recently received some training in evaluation, and was required to test some innovative evaluation tools. Global assessment scaling was presented to the facilitator in the form of a case study from the field of social research. It was presented as a cheap, easy and quick method to assess group performance. The audience of the evaluation was internal to the department.

Method of evaluation

A group of 5 extension workers from the project were selected. Participants were deliberately chosen to represent a range of views towards the project, and were from a range of geographical locations. The central question posed was "are groups providing tangible benefits?" Prior to the first meeting, each of the group members drafted a scale with 5 levels that described successful to non-functional groups. At the first workshop these 5 individual scales were amalgamated to present the views of the whole group.

The scale was then tested with a group of 15 extension workers. The facilitator described a hypothetical group and each of the participants was asked to score the group using the scale that had been previously created. It was found that scores were remarkably close. It was felt that this scale adequately described the various stages through which the community groups were moving. The closeness of the scoring gave the group confidence that an 'outsider' could join the team and use the scale with reasonably consistent results over a period of time.

Table 6 Global assessment scale describing community groups	
Score	***Description***
5	*Most members of the community are contributing to the group and recognise they play an integral part in achieving holistic, long term and agreed community objectives. The group has its own identity and strives for excellence. They are able to identify and implement innovative solutions to problems with little or no government support. Members are willing to accept leadership, responsibilities and different roles. All members are implementing on-ground works and attending regular meetings. The group is exceeding salinity tree and pasture establishment targets and will be able to halt salinity within 30 years.*
4	*Most members of the community have an interest in the group and are working towards a shared long-term strategy. Most members have a holistic and regional vision, but others are still grappling with the concept. All activities are planned carefully by the group and attract significant interest. Government specialists may be invited to provide technical advice. There is a strong committee commitment and other sub-committees are completing specialised roles. The group is meeting salinity targets every year and will significantly slow the spread of salinity in the next 30 years.*
3	*The group shares common medium term goals and is developing a team culture and cohesion. There is a commitment from about 40% of the community to attend meetings regularly and complete on-ground works. New members are encouraged and there is an effort to conduct interesting meetings and activities. Government agencies assist with technical advice and organising activities at the group's initiation. Salinity targets may or may not be met, depending on economic conditions, but there is a significant amount of on-ground works completed each year.*
2	*The group looks to government to set directions and activities. A small group of dedicated members have held leadership roles for long periods and are experiencing 'burn-out'. The group may compete with other organisations for membership, or members may consist of people with specific agendas. There is no long term planning to assist direction setting and goals are strictly short term and self-centred. On-ground works are completed by a small, dedicated core through government funding. Salinity targets are not being met although small areas of salinity may be mediated with time.*
1	*The group is totally dependant on government for funding, support and leadership. There is a reluctance of members to assume any leadership roles or responsibilities, and there is apathy towards attracting new members. Meetings are irregular with few core members present, or meetings are non-existent. There are no agreed goals, and members may not share common problems to bring them closer together. There is little or no evidence of on-ground works occurring. The salinity problem will continue growing.*

Use of findings

The process also led to some serious questioning of assumptions. Some groups, although rating reasonably well, were not actually conducting many on-ground activities to counter salinity. The opposite was also true; some groups given lower scores in terms of group performance had actually completed considerable on-ground activities. So, the assumption that higher rated groups were better at implementing on-ground measures was questioned. It also became evident that the relationship between the degree of facilitation support in terms of facilitation and the score level of group performance was not clear. These things together led to a change in the project approach.

To date the scale has been used several times to understand what stage the groups have reached in several of the department's projects. It has been found that the process of coming to a consensus on the scale can spread clarity in terms of 'what a good group looks like' and the overall aim of group involvement. It can probe the expectations of the project staff and the various perceptions of success.

Further use of Global Assessment Scaling

In addition to project staff using this method to evaluate the progress of a group against a scale, it can also be used by community groups themselves to monitor their own progress. It has been found in other fields that what a group might consider to be successful may vary considerably from what a project staff might deem to be successful. For example, a community group that is good at acquiring government support and is facilitated externally may be given a high score by the group, but given a low score by a project team because they are dependent on government funding and external facilitation.

Source: Personal communication with Cam Nicholson, Victoria

9.2. PROCESS EVALUATION THAT IS PARTICULARLY RESPONSIVE TO NEEDS OF PROGRAM CLIENTS (FARMERS)

In this approach, evaluations are designed to be responsive to the needs of those directly affected. Responsive evaluation is based on the view that those with a direct vested interest in program delivery should control any evaluation agenda associated with the program. Two major models of responsive evaluation are **naturalistic** inquiry, and **action research**.

Naturalistic inquiry

Stake (1967) states that naturalistic inquiry is evaluation which:

- attempts to arrive at naturalistic generalisations on the part of the audience
- is aimed at non-technical audiences like teachers or the public at large
- uses ordinary language, which is based on informal everyday reasoning and
- makes extensive use of arguments which attempt to establish the 'structure of reality.'

There are various models of evaluation that fit into this category including Patton's (1997) **Utilisation Focused Evaluation** and Guba and Lincoln's (1981) **Fourth Generation Evaluation**. Case Study 11 illustrates a process evaluation that was oriented to be responsive to the needs of the clients. This evaluation could be described as naturalistic inquiry in that it did not try to change what the participants do or feel, but instead focused on what the participants are experiencing (Patton 1997). In naturalistic inquiry there is a focus on the use of metaphor, analogy, informal (but valid) inference, vividness of description, reasons-explanations, multiple perspectives and tacit knowledge. Naturalistic inquiry employs qualitative methods (Scriven 1991).

Case Study 11
Improving the usefulness of the Pigpulse Report, an evaluation

Evaluation	*Soft* ◄—●—► *Hard*				
Form	*Formative* ●		*Summative*		
Owen's form	*F5*	*F4*	*F3* ●	*F2*	*F1*
Data collection	*Qualitative* ●		*Both*	*Quantitative*	
Stage of program	*Forming*	*Developing* ●		*Settled*	*Finished*
Program planning level	*Regional* ●		*State*	*National*	

The program

The Pigpulse Report was developed to improve the analysis of piggery data, to help decisions about interventions in piggery management. The Pigpulse project is a service provided to producers involved in the Sowtel-Pigpulse bureau operated under the project Pig Enterprise Management Services (PEMS). To generate the Pigpulse Report, a farm's production figures and some financial records are processed using a computer package. Results of both individual farm productivity and state and district averages are sent back to farmers quarterly. The report lists how the farm production has deviated statistically from these averages. The report monitors traits such as pre-weaning mortality, average birth weights, and weaner weights.

In addition to the report, producers participate in discussion groups in a number of districts. These groups discuss issues raised by the reports, and the groups are encouraged to determine some of the topics. This evaluation concerns one aspect of the PEMS project; the Pigpulse project. This project was in the process of being developed at the time of evaluation.

Goal of evaluation

To find out if the development of the Pigpulse report was useful to producers in supporting decisions.

Form of evaluation

This evaluation best fits Owen's Form 3, process evaluation. The evaluation information was aimed at improvement and development of an existing activity. It involved some assessment of how the activity is reaching its goals, but concentrated on how the activity can be improved and made more effective. There was an opportunity to make changes in the software development as well as in the extension.

Actors

The evaluator was internal to the program, as was the primary audience for the evaluation.

Method of evaluation

The pre-test *The evaluation was conducted by pre-testing an interview guide during a semi-structured interview with three producer clients. This guide contained open-ended questions and prompts that were developed from the key question and hypothesis. The interviews were conducted on the producers' farms at a time to suit them. The interviews were arranged by telephone, the producers being informed of the interview content and purpose. The interview guide was used while the producer was asked to look at their last report and to recall how they used the report and what else was considered in making decisions in their piggery management. A copy of the Pigpulse report was provided as interviewees might otherwise say what they think they should consider instead of what they actually did consider.*

During the interview, notes were written by the interviewer and were checked by the producers and modified to include their corrections. The results were coded according to a series of statements derived from the hypothesis. Examples of the statements in the coding frame were:

- *whether producers read and thought through the Pigpulse report*
- *whether trait type was considered*
- *whether the magnitude of change was considered*
- *whether targets were compared with the new level*
- *whether the cause of change was considered*
- *whether the effect on profit was important*
- *whether producers changed their practice.*

As a result of the pre-test, the interview guide was altered before the focused group discussions (FGD) took place. FGDs were chosen rather than questionnaires as the producers were feeling over-surveyed.

Use of findings
The results of the interviews were related to the level of Bennett's hierarchy that had been affected by the program. The report was an evaluation in process and did not include findings of the focused group discussions. However, it suggested that these results would be expressed qualitatively.

Source: Prepared from a report by Alison Spencer, Rural Extension Centre, U. Queensland, Gatton

Action research

Action research, as its name implies, pursues both action and research outcomes. It is most commonly done with high levels of participation. Its purpose is to bring about planned change in such a way that there is also a contribution to understanding of the system which is being changed, and the methods which are used. The term 'action research' is applied to a confusing array of activities.

It could be argued that all formative evaluation is action research, in that it is research that leads to action (the refinement of a program). Hart and Bond (1995) selected seven criteria to distinguish action research from other methodologies, i.e. it:

- is educative
- deals with individuals as member of social groups
- is problem-focused, context-specific and future-oriented
- involves a change intervention
- aims at improvement and involvement
- involves a cyclical process in which research, action and evaluation are interlinked
- is founded on a research relationship in which those involved are participants in the change process.

Some forms of participatory evaluation or empowerment evaluation (Fetterman 1996) could also be described as action research. In this study, none of the evaluation reports reviewed could be clearly described as action research. However, concepts of action research are increasingly used for evaluation purposes in agricultural extension in Australia. It is reasonable to suggest that as action research is not generally labelled as 'evaluation' it may have slipped through the net of the Review. However, it is clear that action research does involve iterative cycles of action and evaluation.

10. FORM 4: EVALUATION FOR DESIGN CLARIFICATION

Evaluation for design clarification is generally carried out when the program is at the developmental stage. Examples of methods used in this form are **evaluability assessment** (Wholey 1983, Smith 89), **program logic** (see below).

10.1. PROGRAM LOGIC

Program logic is generally drawn as a visual diagram of the events of a program which lead to the ultimate outcome. It enables the participants to question the cause and effect assumptions between the various steps. It is best used as a team approach and is noted for enabling staff to come to consensus about the realistic outcomes and goals of a program. Ideally, program logic would be mapped out by the project team before the implementation and modified and referred to throughout the life of the project. However, in many cases program logic is conducted later on in the project life to help in the evaluation process, or to bring a program back on track.

Case Study 12 is presented here as it is an example of the use of program logic in evaluation for the purpose of clarifying the design of a program once it has commenced and there is dissatisfaction about its progress or original design.

Case Study 12 ***A workshop to refine the program logic for Coranagmite Salinity Program***					
Evaluation	*Soft* ◄——●——► *Hard*				
Form	*Formative* ●	*Summative*			
Owen's form	*F5*	*F4* ●	*F3*	*F2*	*F1*
Data collection	*Qualitative* ●	*Both*	*Quantitative*		
Stage of program	*Forming*	*Developing*	*Settled* ●	*Finished*	
Program planning level	*Regional* ●	*State*	*National*		

The Program
The Victorian DNRE (Department of Natural Resources and Environment) decided to have a rigorous look at one of its programs directed against the problem of dryland salinity. This program had been running for about 5 years, and was considered to be at a settled stage. The specific project to be evaluated was the Coranagmite salinity program.

Goal of the evaluation
To gain consensus as to the project aims and to improve the implementation of the project.

Form of evaluation
This evaluation best fits Owen's Form 4, evaluation for design clarification.

Actors
The group that requested the workshop was the DNRE staff responsible for this program, i.e. an internal audience. Other stakeholders included the steering committee and an implementation group, and landholders. These groups were not included at this stage. The facilitators were from the DNRE, and had recently received training in program logic.

Method of evaluation
Two workshop sessions were carried out over 1½ days. 8 people attended and the mode of inquiry was informal discussion and brainstorming. The steps taken are outlined below:

Step 1. Defining the problem: *In the first session the group set about defining the problem. The facilitators encouraged participants to think critically. At this stage it became apparent that there was no common understanding of the underling problem. The participants were encouraged to describe the 'rich picture', the political nuances, and all the stakeholders. Once they had agreed on what the problem really was, they moved on to discussing what outcome would be achieved if this problem were solved. When a consensus had been achieved as to realistic outcomes, they moved on to the next stage.*

Step 2. Mapping the first layer of the model. *Participants were asked by the facilitator to try to forget that they were involved in the program in reality and to imagine that they were in the planning phase of a new program. They were then asked to think about what was needed in place in the hypothetical program for the outcome to be achieved. This formed the first level of the hierarchy in the program logic map.*

Step 3. Mapping out subsequent layers of the model going downwards from the outcome: *The participants were asked how these necessities were to be achieved, and thus the map of program logic was created moving downwards from the desired outcome to activities necessary to achieve these. At each level of the map the facilitators would double check with the participants, by asking questions such as: "are you really sure?" and "if you do this and this, will that really lead to the outcome?" At each step of the model the cause-effect relationship was explored. The facilitators constantly questioned the participants with regards to any assumptions they had made concerning the links between cause and effect.*

Step 4. Back into the real program*: Participants were asked to return to the reality of the current program and compare the conceptual model with that of current activities and state of play of the real program. It became obvious that certain things were not being done in the dryland salinity program that should have been done and also that certain inappropriate activities were being carried.*

Step 5. Considering evaluation*: This stage was to consider where evaluation should take place. The facilitators suggested that the links between cause and effect that involved 'leaps of faith' should be evaluated. Participants were asked to identify these weak links in the model. Once the locations or links had been established, the participants were asked to consider which data would answer their questions and what evidence was available to check that one link in the program led to the next intermediate outcome.*

Presentation/utilisation of findings
The results of these workshops were not formally recorded. However, 5 days after the workshop ended, the DRNE members of the dryland salinity program were involved in an independently- organised workshop with the whole dryland salinity program team of Coranagmite catchment area. The participants of the previous program logic workshop rescheduled the whole workshop to incorporate what they had learned in the previous workshop and to share it with the rest of the team. The program design was significantly altered to address the issues raised in the first workshop. This constitutes a high degree of utilisation, according to Patton's framework (1997).

Source: David Beckinsale (1997) Department Natural Resources and Environment, Melbourne

11. FORM 5: EVALUATION FOR PROGRAM DEVELOPMENT

Form 5, evaluation for program development, is for the synthesis of a new program. In this case, the program will not yet have been implemented and will still be at the planning stage. The focus is on the context of the program environment. Examples of models and methods include needs assessment, economic evaluation, review of best practice, research synthesis, exploratory RRA (rapid rural appraisal) or PRA (participatory rural appraisal), and the Delphi technique.

11.1. ECONOMIC EVALUATION FOR PROGRAM SYNTHESIS

Some organisations conduct economic exercises to decide whether a proposed project is viable or not. Cost–benefit analysis is one method that can be used (see Section 7.3).

11.2. NEEDS ASSESSMENT

According to Owen (1993) needs assessment is the best known and most acknowledged approach to evaluation for program development. However, Scriven (1991) warns that the term 'needs assessment' (or needs analysis/needs sensing) has drifted from literal meaning into jargon status in which it refers to any study of the needs, wants, market preferences, values or ideas that might be relevant to a program. True needs are considerably harder to establish than felt wants because true needs are often unknown to those who have them, and may be contrary to what they want, as in the case of a boy who needs a certain diet and wants an entirely different one (Scriven 1991). Owen (1993) explains that an entity can be defined as 'in need' if there is a difference between the actual situation in which it exists and the desired situation at that time. The needs assessments should be concerned with establishing:

1. the desired or ideal state of affairs
2. the actual state of affairs
3. discrepancies between 1 and 2
4. reasons for the discrepancies or needs
5. what needs should be given priority for action through program.

While conducting the Review we found few examples of needs assessment. In extension, needs assessment is often considered as research rather than evaluation and may have slipped through the net of the study. However, Case Study 13 is an example of needs assessment which uses participatory rural appraisal (PRA).

11.3. RAPID RURAL APPRAISAL

Rapid rural appraisal (RRA) and participatory rural appraisal (PRA) are used extensively in social and natural resource projects in developing countries. Rapid rural appraisal emerged in the late 1970s as a reaction to the general dissatisfaction with the biases inherent in the way much of rural development research tended to operate. These biases included:

- spatial consideration – researchers tended to visit places that were easy to get to
- gender and poverty – researchers tended to talk to men rather than women
- seasonal – that researchers tended to visit during the cool dry weather rather than hot and wet periods.

RRA was also a reaction against the expense and frequent inaccuracy of the conventional process of questionnaire survey and recognition of the value of indigenous technical knowledge (IT). RRA practitioners asked the question, "whose knowledge counts"? and have attempted to shift the balance from an emphasis on information and knowledge that outsiders know and accept, to that which farmers know and accept. It developed largely from farming systems research, but was also influenced by:

- Activist participatory research inspired by Paulo Friere (1972) who was a proponent of the 'popular education' movement in Latin America. This was largely about dialogue and joint research to enhance people's awareness and to empower them to take action. Its key contribution to RRA was its recognition that poor people were creative and capable and should be empowered, while outsiders have a role as catalysts and facilitators (Pretty *et al.,* 1996)
- Agroecosytems analysis – developed by Gordon Conway and colleagues (Conway 1985), this approach draws on systems thinking and ecological activity. Among its major contributions to current approaches are its use of transects, informal mapping and diagramming
- Applied anthropology emphasised the benefits of unhurried participant observation and conversations and the importance of behaviour and rapport.

A multidisciplinary team staying in the community for a period of around six days generally carries out RRA. The techniques used in RRA include semi-structured interviews, diagramming, stories, histories, diagrams, conceptual representations, maps, transects, seasonal calenders, time lines, flow diagrams, decision trees, venn diagrams, pie charts and more. An important component of RRA is the reliance upon 'triangulation': this involves crosschecking information by talking to different stakeholders, using different methods (interviews, diagrams, and observations) and having a multidisciplinary team (men, women, insiders and outsiders). Triangulation is carried out to assure the validity of the findings by strengthening/optimising the sources of information (Pretty *et al*.1995).

11.4 Participatory Rural Appraisal

In the 1980s to 1990s in developing countries' programs, the RRA began to be replaced with PRA. The change from RRA to PRA was not just one of terminology, but of philosophy and objectives. RRA and PRA share common methodology, but the aims to which they are put differ. PRA is basically about facilitating and contributing to the empowerment of local people to take control over their own appraisal of problems and opportunities, and to select appropriate action. The concept of 'handing over the stick' was central to this shift (Chambers 1994).

PRA/RRA in Australia

In the last decade modified versions of RRA and PRA have begun to be used in industrialised country contexts. Generally, the reliance is more upon multidisciplinary teams, and a strong use of semi-structured interviews. Less emphasis is placed upon diagrams and pictures. In the context of evaluation, RRA/PRA can be used to develop new programs, for process evaluation and for impact evaluation. It is most appropriate when a high degree of client feedback is required in the evaluation. In 'evaluation talk' it would be best described by naturalistic inquiry. Case Study 13 is an example of an RRA that was carried out in order to contribute to new programs that were being planned.

Case Study 13
Understanding farmer-decision making on land use

Evaluation	*Soft* ◄—●—► *Hard*				
Form	*Formative* ●		*Summative*		
Owen's form	*F5* ●	*F4*	*F3*	*F2*	*F1*
Data collection	*Qualitative*		*Both* ●		*Quantitative*
Stage of program	*Forming* ●	*Developing*	*Settled*		*Finished*
Program planning level	*Regional* ●		*State*		*National*

The program
The Department of Primary Industries and National Soil Conservation Program (NSCP) is funding a major project to hasten the exchange of knowledge and management skills amongst southern Queensland farmers on fallow management and viable farming systems. The VFSG is a DPI project team integrating extension and research personnel across southern Queensland. The project commenced in January 1991 and will run for 5 years.

To ensure that VSDG programs and products would be relevant to users, and that exchange of respective information would be efficiently conducted, farmers attitudes, knowledge and needs on farming system were researched. This was part of a research project set up because of concerns over excessive soil erosion, rainfall run-off, declining soil nitrogen fertility and soil structure, sub-optimal use of herbicides and fertilisers, and insufficient rotation of crops and pastures. The research was undertaken to ensure that consequent VFSG programs and products would be relevant to users and exchange of respective information would be efficiently conducted.

Goal of evaluation
Was to:

- *develop an understanding of farming systems and farmer attitudes towards farming systems in southern Queensland*
- *gain an overview of the important issues as seen by the people involved in the management of these relevant farming systems*
- *determine which issues required further investigation and who possessed knowledge that would assist in these activities*
- *provide a base of common understanding for further market research and action.*

The concept of triangulation was used in the methodology. Triangulation involves the use of different methods to study the same problem or situation and provides a way of crosschecking data.

Form of evaluation
This evaluation best fits Owen's Form 5, evaluation for program development. In evaluation talk it can be described as a needs-assessment using participatory methodology.

Actors
The evaluation was carried out internally to an internal audience.

Method of evaluation
The research consisted of two phases:

- *a modification of Rapid Rural Appraisal*
- *focused group discussions to crosscheck issues raised in the RRA, investigate issues in greater depth and to design a quantitative survey for future progress and impact evaluation.*

The rapid rural appraisal consisted of 100 face-to-face semi-structured interviews carried out by 2 multi-disciplinary teams. Each team consisted of 4 interview pairs with one member from the VFSG and one from outside. In this context, RRA was used to produce a wide overview of possible aspects, people and interactions and to obtain an impression of which issues were likely to be the important. RRA is meant to be heuristic, in that it will tell you what is there. The aims of the exercise were achieved by: reaching a wide cross-section of the relevant community, adjusting the definition of what was relevant during the exercise, and not excluding any related material from the report. The interviews were conducted in pairs (or small groups) with different backgrounds.

The focused group discussions were arranged into 5 districts and were segregated with regard to their tillage practice. Participants had to be located within 60 minutes travel time of the focus groups venues and those farmers involved in the RRA were not involved. Farmers who met the above criteria were selected randomly from DPI and Grainco mailing lists. The prospective respondents were contacted by telephone and invited to attend the focused group discussions. Moderators were used in the discussions to focus the groups in a non-directional manner. Interviews were taped to assist analysis. A typist transcribed the transcripts from the FGDs and a computer package called ethnograph *was used to analyse the data from the focused group discussions and search for patterns and themes.*

Use of findings
Implementation of VSG sub-programs involving many of the recommendations from this research has commenced.

Source: Taken from 2 internal documents: (1) N.A Hamilton (DPI Toowoomba) and (2) David Blacket and Gus Hamilton, DPI ,Goodndiwindi Queensland (1992)

12. FRAMEWORKS FOR PLANNING PROGRAM EVALUATION

So far in the Review case studies have been categorised with regards to the ultimate purpose of the evaluation, which we have termed 'form'. However, in reality the evaluation of a program can often employ more than one form. Figure 10 shows an evaluation strategy for a hypothetical program, which includes four different Forms of evaluation. In this Section we take a more 'program-focused' view of evaluation and review some of the frameworks that can be used to form an integrated evaluation plan for a program.

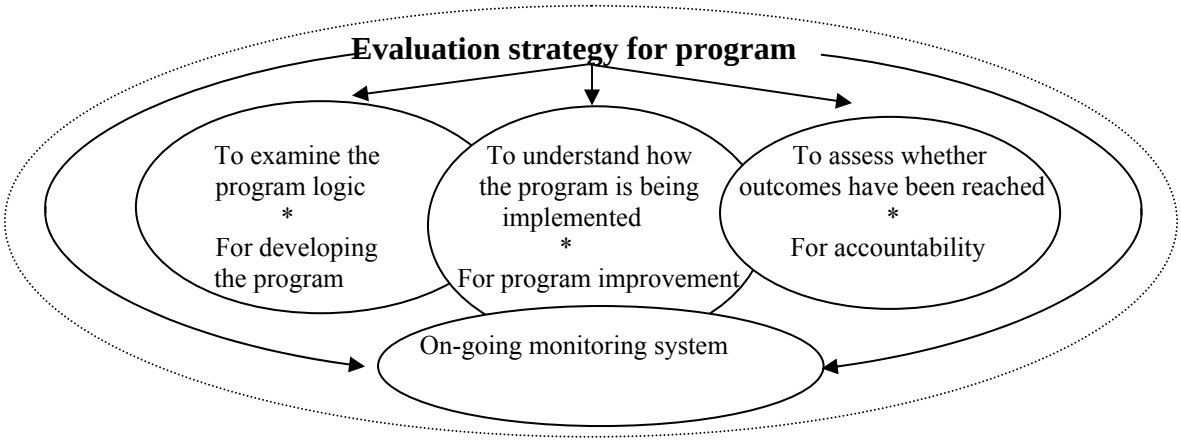


Figure 10 Evaluation strategy for a hypothetical program that employs multiple forms

All evaluations have a conceptual framework at their heart. In some well-resourced evaluations with long 'lead-in' times such frameworks are usually explicitly identified. In contrast, in many smaller-scale exercises the framework is implicit and may be seen as 'common sense'. According to Hulme (1997) there are three main elements to a conceptual framework for an evaluation:

- the model of the impact chain that the study is to examine
- the specification of the units or levels at which impacts are assessed
- the specification of the types of impact that are to be assessed.

In developing an overall framework for evaluation the key role for the evaluator is to assist program staff to set out their hierarchical objectives and to identify those goals that can be used to develop indicators of progress towards desired outcomes. Several logical frameworks are currently used in agricultural extension to help plan an evaluation strategy, including:

- Bennett's hierarchy – a seven-level ladder that conceptualises hierarchical levels of program outcomes
- the logical framework – a hierarchy of objective statements regarding the goal, purpose, outputs, and inputs
- the Snyder model – a soft-systems approach to evaluation
- the three-rings approach – differs from the other frameworks in that it recognises that other agencies and programs may also have an impact on the target group.

Program logic models and the 'team up' approach

All the above frameworks can help practitioners clarify what the desired outcomes are and identify the steps that will lead to these desired outcomes. A common feature of these approaches is the way that they are utilised most effectively. If one team member is left to draw these logic models alone, then much may be lost. In some organisations an approach referred to as 'team up' is adapted where the frameworks are drawn as a team. This can lead to considerable team building, common understanding of goals and questioning of assumptions.

12.1. BENNETT'S HIERARCHY

Bennett's hierarchy is one of the most commonly quoted frameworks in agricultural extension literature and is presented in a popular extension text by Van den Ban and Hawkins (1996). It can be a valuable guide in planning extension programs and developing an evaluation strategy. Bennett describes a chain of events assumed to characterise most programs in extension. He uses this chain of events to depict a hierarchy of objectives and evidence for program evaluation. The hierarchy presents a framework for a goal-based approach to evaluation. The essence of this is that 'the more nearly the objectives of a program are reached, the more positive the judgement of the program'. Bennett's hierarchy consists of seven levels of objectives and evaluative evidence. Table 7 lists eight levels and gives examples of goals for each level, the additional eighth level was added by Van den Ban and Hawkins (1996).

Table 7 Bennett's hierarchy with examples

Level	Description	Examples of objectives
Level 8	Consequences for society	As a consequence of production increase, export targets are reached and the national economy is strengthened.
Level 7	Consequences for the target group	As a consequence of the new practice, production has increased by 25%, contributing to an increased farm income of the target group, and a decreasing trend for target farmers to move out of the area.
Level 6	Behavioural changes in the target group: Direct evidence Indirect evidence, product quality – as evidence of change in behavour)	2 years later, it is found that 50 % of the community have adopted new practices, or prime lamb production rose by 12 percent
Level 5	Change in: knowledge attitude skills motivation and group norms (Proxy indicators of behaviour change)	Differences in the findings of surveys that were conducted before and after the intervention indicate that the target farmers have gained knowledge and skills
Level 4	The farmer's opinion about extension activities	In a one-page survey completed at the end of the workshop 80% of the participants were satisfied with the program.
Level 3	Farmer participation in extension activities (participation monitoring)	400 farmers attended the meetings, 500 listened to the broadcast, and more than 35% of the attendees were women.
Level 2	Implementation of the program by extension agents (activities monitoring)	20 workshops held, 2,000 pamphlets published and distributed, 5 radio programs conducted to advertise the event.
Level 1	Programming of the extension activities (inputs monitoring)	7 staff will be trained, and an equivalent of 7 full time wages spent on the project
Source : Adapted from Bennett (1977)		

Bennett (1977) suggests a series of guidelines to help decide which levels of the hierarchy should be examined in an evaluation. The first three guidelines are summarised below:

(a) Evidence of program impact becomes stronger as the hierarchy is ascended. Evidence at the two lowest levels provides little or no measure of the extent to which clientele benefit from the program. Level 3 merely provides one way of measuring possible opportunity for education to occur, while level 4 can provide somewhat better confirmation of whether given activities have been beneficial, but is less satisfactory than level 5 (KASA). Changes in KASA are considered to be merely stepping-stones to indicate the adoption of more desirable patterns of behaviour. Level 5, (evidence of practice change) is desirable when program objectives include changing practice. Finally assessing practice change is usually quite apart from assessing accomplishment of ultimate program objectives. Extension is often held accountable for the extent to which it is contributing to solution or checking of overall problems. Therefore an ideal evaluation of impact would probably be in terms of whether desired end results are achieved plus any significant side effects.

(b) The difficulty and cost of obtaining evidence of program accomplishment generally increases as the hierarchy is ascended. This is due to: an increasing time–lag needed to asses whether the objectives have been reached; increasing difficulty in attributing effects to the program being evaluated; increasing difficulty in setting precise objectives and increasing expense.

(c) Evaluation is strengthened by assessing extension programs at several levels of the hierarchy, including the inputs level.

12.2. THE LOGICAL FRAMEWORK

The logical framework (matrix) is used extensively by Australian and other agencies working in overseas agricultural projects. As many Australians have used it overseas, it is surprising that the technique is hardly ever used within Australia in agricultural extension. Logframes are very helpful in identifying the essential elements of a project and the linkages between these in a logical, concise and objective manner. Thus it can introduce order and discipline into the project design process and helps to place the project in the larger context, e.g. of a regional plan. The framework also serves as a tool for logically identifying inputs, assumptions and success and indicators for monitoring progress and evaluating performance. Thus the Logframe is both a planner's tool for setting out the design elements of a project, and a tool for managers to review and report on progress. A general description of the logical framework is provided below, followed by an example of its use in Case Study 14.

A logframe comprises a hierarchy of objective statements regarding the **Goal**, **Purpose**, **Outputs**, **Inputs** (which include activities). These objectives are listed in column 1 of a Logframe (see blank frame below). There are causal linkages between each level of the hierarchy. Inputs are used with activities to produce outputs, outputs are expected to achieve project purposes; the achievement of purposes is expected to contribute to the overall goal. The basic premise is that the achievement of each level is needed for the achievement of the next higher level. The progress from one level to the next is conditional on the validity of the stated assumptions, which are stated in the right hand column. Tabel 7 shows the questions that need to be answered when filling in each cell of a logframe. Definitions of the terms used in the logframe are:

- **Goal:** A statement of overall development aim (or vision) of the project. Usually there is only one goal, which is very broad and often related to a wider national or regional objective. The project would not be expected to achieve this alone, but to contribute towards this goal
- **Purpose:** The primary reason for the project – related to the problems that a project is designed to address, development expected during the project or the situation envisaged at its completion. A project may have more than one purpose; each must be stated explicitly
- **Outputs:** Project outputs are the direct, identifiable and measurable results expected from the provision of inputs. They are the preconditions for the achievement of project purpose
- **Inputs**: Project inputs are the resources made available, which together with the activities, allow achievement of outputs. Inputs may be people, equipment, finance.

- **Achievement:** Project activities are the discrete tasks undertaken using resource inputs to achieve defined outputs. Observation of execution of activities forms a basis for monitoring performance.

An integral aspect of the logframe is that **assumptions** are stated about the effect of the external environment on achievement of the goal, purpose, outputs of the project. They indicate that there are elements of uncertainly. In designing a project, planners must address both the vertical and horizontal logic of the logframe. The vertical logic is based on the principle of causality between levels, while the horizontal logic identifies and states the means by which the project will produce results at the various levels. **Verifiable indicators** are stated in the second column, and method of collecting the information is recorded in the third column. In Table 8 the questions in each square provide brief instructions on how to complete a lograme matrix. An example of a complete Logframe is presented in Case Study 14.

Table 8 The logframe matrix: questions to be answered when filling in each cell of a logframe.

	Narrative summary	Measurable indicators	Means of verification	Assumptions
Broader goal	What wider issue will the project help to resolve (e.g. a national objective)?	What are the indicators of goal achievement?	What are the sources of information? What methods are to be used to obtain it?	**(Purpose to goal)** What assumptions/external factors must be true if the purposes are to help reach the Goal?
Purpose [Primary reasons for the project]	What immediate changes/ benefits will the project bring about for the target population or area?	What are the measures to judge the project's immediate effects, benefits and losses?	What are the sources of information? What methods are to be used for obtaining it?	**(Outputs to purpose)** What external factors must be true if outputs are to achieve the Purpose?
Outputs	What direct outputs will the project produce so it can achieve its purpose?	What kind and quantity of outputs are planned, and by when will they be produced?	What are the sources of information? What methods are to be used for obtaining it?	**(Inputs to output)** What external factors must be realised to produce the planned outputs in time?
Activities	What activities must be undertaken to produce the outputs? When must these activities take place?	**Inputs/ resources** What materials, equipment or services are required, at what cost, over what period?	What are the sources of information on inputs? What methods are to be used to obtain it?	**(Initial assumptions)** What actions outside the control of the donor are necessary to begin the project?
Source: Modified from AIDAB (1991) and Farrington and Nelson (1997)				

Case Study 14
A logframe for proposed project to support Landcare group activities in the Wimmera River Catchment.

Evaluation	*Soft* ◄——●——► *Hard*				
Form	*Formative* ●			*Summative*	
Owen's form	*F5* ●	*F4*	*F3*	*F2* ●	*F1*
Data collection	*Qualitative*		*Both* ●		*Quantitative*
Stage of program	*Forming* ●	*Developing*		*Settled*	*Finished*
Program planning level	*Regional* ●		*State*		*National*

The program

The project was a support program for Landcare group activities amongst 6 Landcare groups and was out-sourced by a natural resource agency (tenders called for in a public advertisement). The contractors were required to estimate expenditure for only an initial 12-month period, but to place this in the context of a longer-term program and the overall goal of the agency.

How frameworks related to the evaluation

A logframe was used to develop and present a proposal (tender) for the Landcare support project. One aim in adopting a logframe was to use it as a means of rapidly seeking and documenting opinions of members of the planning team. Another goal was to present the project proposal in a very concise manner to the funding agency. In addition to its use in project design, the logframe was intended to provide a framework for monitoring and evaluation of progress and problems during the life of the project.

Description of the framework

This logframe (next page) was constructed with a small, diverse team of people, most of whom had never used a logframe previously. The process started by seeking agreement on the broad goal of the agency (top left of frame), and then the 'purposes' under that. The assumptions made for the purposes to achieve the goal were discussed and stated. Then the outcomes need to achieve the purpose were listed, followed again by the assumptions. Moving down the frame, the activities needed to achieve the outputs and the assumptions were stated, and finally the initial inputs in terms of staff, time and dollars for the first 6 months of the project.

After a few days, the team met again to complete the two middle columns on indicators and means of verification. This led to some modification in the purposes and outcomes etc., but within a total of a few hours the team had completed the matrix. The frame was modified and improved later, on the basis of comment from other stakeholders who were not at the discussions.

Use of findings

In developing the logframe, surprising differences in views were identified amongst team members, which were resolved in the process of completing the task. The logframe was submitted as part of a tender and was to form the basis for monitoring and evaluation of the project itself.

Experience elsewhere shows that logframes completed at the project planning stage can be very valuable for management purposes throughout project life, give useful direction to staff, and help to identify false assumptions made earlier so that changes can be made soon enough to avoid waste of resources. The logframe is shown on the following page.

Source: John Petheram, Longerenong College, University of Melbourne, Horsham

LOG FRAME for a proposed project for support of Landcare groups in a catchment area

NARRATIVE SUMMARY	VERIFIABLE INDICATORS	MEANS OF VERIFICATION	ASSUMPTIONS
<u>GOAL</u> Sustainable farming in the northeastern region of the Wimmera River catchment	1. Reduced levels of land degradation on farms 2. Improved economic viability of farms 3. Improved levels of salinity & water quality in catchment	1. Monitoring of soil infiltration rates & aggregate stability (10 sites) 2. Data on debt levels from ABS and local rural counsellors 3. Data from regional catchment monitoring systems	<u>Assumptions (Purpose -> Goal)</u> 1. Adoption of improved farming & monitoring practices leads to: • reduced land degradation • improved viability of farms 2. Farmer identification of degradation issues leads to acceptance of responsibilities to improve catchment 3. Funds will allow necessary works to tackle degradation issues
<u>PURPOSES (1-3 years)</u> 1. Adoption by farmers of practices that are economically and ecologically sustainable 2. Farmer adoption of methods of monitoring condition of land and other resources 3. Identification by farmer groups of principles & action needed in improving catchment condition 4. Financial support obtained for efforts by community groups to undertake works/ activities aimed to improve catchment condition	• Change in farmer practices related to resource management • Establishment of systems for monitoring condition of farm natural resources • Community success in gaining finance for conservation activities • Community satisfaction with progress and service provided	• Rapid appraisal survey of change in farmer practices after 2 years • Rapid appraisal survey of farmer involvement in monitoring after 2 years. Review of monitoring data • No. and value of grants received in 1996 & 1997 • Informal study of community group satisfaction with progress/service	<u>Assumptions (Output Purpose)</u> 1. Farmer training and the conduct of trials on farms by groups will lead to adoption of economically and ecologically sound practices 2. Training and trials will lead to farmer adoption of monitoring methods
<u>OUTPUTS (first 12 months)</u> 1. Farmers undertake training in land management issues/monitoring methods important to their needs 2. Groups plan and conduct farm trials of methods to improve viability and sustainability of farm practices & to monitor land degradation 3. Group applications made for funding to conduct farm trials, training, other activities	1. Training courses and specialist information sessions 2. Farm trials planned with specialists to test ideas for improvement. 3. Groups prepare funding submissions for trials and other group activities (e.g. tours)	• Three short courses planned for 1996/7 • 4 information sessions/ seminars planned for 1996/97 • 4 farm trials planned for 1996/7 • 4 funding applications prepared and submitted in 1996	<u>Assumptions (Activity Output)</u> 1. Farmers will accept need for information & training on land issues 2. Farmers will accept need for farm trials to test & demonstrate methods 3. Solutions exist which farmers are will verify as means of overcoming land management problems
<u>ACTIVITIES (first 6 months)</u> 1. Team members visit all groups & major land management issues & group works to date 2. Team listen to group's aims/priorities/plans 3. Assist groups to clarify aims & make plans 4. Run specialist issue sessions with groups 5. Promotion of group activities in media 6. Assist funding applications for group needs	<u>Inputs/resources (first 6 months)</u> • 30 farm/client visits • 10 group activities • 4 major events ($1200) • 6 media promotions ($1200) • process 6 LPIS applications • 8000 km (@ $0.50) = $4000 • audio/photography costs $2000 • specialist fees $5000		<u>Initial assumptions</u> Serious drought or crop failure would detract from conditions for successful group activities on cropping farms.

12.3. THE SNYDER APPROACH TO EVALUATION

The Snyder model can be used to guide simultaneous outcome evaluation (equivalent to impact evaluation – Owen's Form 1) and long-cycle process evaluation (equivalent to monitoring and evaluation for program management – Owen's Form 4). During the Review a comprehensive effort was made to find an example of an evaluation study of an agricultural extension program that had used the Snyder approach. Several practitioners reported using elements of this model, or currently setting up evaluation using this model, but no example was found of a completed evaluation using this approach. Because of this a case study of a pilot program to replace imprisonment for offenders is described. Case Study 15 was chosen particularly to illustrate the combination of all three components of the process.

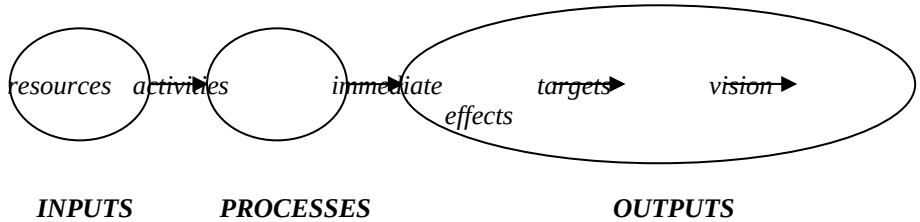
Case Study 15
Evaluation of a program to replace imprisonment using the Snyder model

Evaluation	*Soft* ◄—●—► *Hard*				
Form	*Formative* ●		*Summative* ●		
Owen's form	*F5* ●	*F4*	*F3* ●	*F2* ●	*F1* ●
Data collection	*Qualitative*		*Both* ●		*Quantitative*
Stage of program	*Forming* ●	*Developing* ●		*Settled* ●	*Finished* ●
Program planning level	*Regional* ●		*State*		*National*

The program
The program set out to implement alternatives to imprisonment as a means of sentencing. The evaluation was commissioned to take place near the completion of a trial period. It was to help the government of the day decide how far they would extend the program. This model can be used to monitor and evaluate any program, but is particularly suitable for programs that are participatory in approach.

Approach and methods of evaluation
The Snyder model applies systems theory to evaluation. It consists of a conceptual model and a set of procedures, which allow an integrated participatory evaluation to be conducted. It is used when the program is already running. It allows simultaneous long-cycle outcome and process evaluation: that is, it allows after-the–event evaluation which both assesses goal achievement and reviews activities and processes. The data generated in this approach are generally qualitative. Its main feature is that program outputs are subdivided into 3 categories operating over different time perspectives (see figure below). The program outputs are classified into 3 types:

- *immediate effects of program activities*
- *targets are presented as goals, usually over time spans of months to years (benchmarks)*
- *the vision consisting of the ultimate goals which the unit or project pursues.*



Purpose of evaluation
The central purpose was a summative evaluation of the trial period of the program. Those commissioning the program wanted an answer to the question, "how viable an alternative is this program to imprisonment and other alternatives to imprisonment?" Those directly involved in the program liked the idea of also gathering information that would enable them to improve the program, and build in feedback loops. All 3 components of the Snyder process (process evaluation, outcome evaluation and short-cycle evaluation) were therefore used.

Actors
Much of the data collection and interpretation was done by the director and officers, with the evaluators acting primarily as facilitators. The evaluators wished to involve at least these stakeholders in the evaluation:

- *the staff directly involved in the program*
- *equivalent staff from other programs*
- *a sample of the offenders released from prison onto the program.*

However, an earlier search of local newspapers revealed that the program was electorally unpopular and therefore politically sensitive. Direct involvement was therefore limited to the program director and the officers who supervised the offenders.

Method of evaluation
Part 1. The process evaluation component.
The steps in the procedure are:
- *define the vision*
- *independently define the targets*
- *compare the target to the vision, noting which does not link to the vision, or any elements of the vision not represented in the targets. Adjust targets or vision accordingly*
- *define the activities and use these to identify the actual immediate effects, both intended and unintended*
- *compare the effects and targets, noting any mismatch. Adjust targets accordingly*
- *identify the resources consumed*
- *compare the activities and resources, noting any mismatch. Adjust activities or resources accordingly.*

In terms of time and effort this is the major part of the process. A vision for the program was developed by the participants and checked against legislation, and required administrative procedures. Targets were then defined, and the links between targets and vision identified. This led to modification of both vision and targets. In a similar vein, targets were compared to immediate outcomes. Activities were then examined to check that the most resource-consuming activities were also the activities that contributed most to the vision.

In this case study one component of the vision was that offenders would become law-abiding members of the wider society. A related target was, "low offence rate by people while they were on the program". An activity that contributed to this was regular, partly random, visits by supervisors to offenders. An immediate effect that could be expected to contribute to the vision was that there would be an enhanced supervisor-offender relationship so that offenders would voice concerns and troubles early, when something could be done about it. The associated resources were relationship-building skills on the part of supervisors; time after each visit to think back over the visit and identify what the supervisor could have done to build a better relationship.

Part 2. The outcome evaluation component

Outcome evaluation in this context seeks to evaluate the program in terms of the vision it is pursuing. The procedure is:

- *Note the targets that will later be evaluable. Devise base measures where necessary*
- *Trace targets back to immediate effects, giving particular attention to important but hard-to-evaluate targets. Check if the corresponding immediate effects are evaluable*
- *Trace targets and immediate effects back to activities, giving particular attention to important but hard-to-evaluate targets. Check if the corresponding activities are evaluable*
- *Trace targets, immediate effects and activities back to resources. Give particular attention to important but hard-to-evaluate targets. Check if the corresponding resources are evaluable.*

After understanding the links between resource use, activities and the various levels of outcomes, participants were able to identify performance indicators that they could use to assess program effectiveness. In this instance, 2 sets of indicators were developed. One set, largely quantitative and 'objective', were used to prepare a report to the government. The other set, much of it qualitative, was used by supervisors and director to decide what they could do to improve the program.

Offence rates of participants in this program were compared to those on parole. The participants on the two programs were very similar. The supervision philosophy was very different. The offence rates were spectacularly different. In the official report different offence rates were used to justify a conclusion that the intensive relationship-oriented supervision was an important part of the program. 'Quality of relationship' wasn't something that could be easily used as an official indicator. However, for the supervisors, it was clearly one of the most important ones. It was after thinking about this that the supervisors decided to spend 5 minutes after each supervision visit to reconstruct the visit in their minds, and identify what had worked well and what they might do differently.

Part 3. The short-cycle evaluation component
In this context short-cycle evaluation is a form of evaluation where the people in the program evaluate the actions as they carry them out, and adjust them where necessary. This form of evaluation is more an attitude than anything else, although there are ways of providing encouragement for it to happen. In this evaluation feedback systems were developed to collect the indicators developed in part 2 (above) and apply them to ongoing program development.

Use of findings
The government of the time chose to ignore the recommendations. The supervisors, on the other hand, reported later that the exercise had helped them very much to understand what they were trying to do, and how they might do it better. The experience gained in using this approach to evaluation has generally been that as people come to understand the links from resources to activities to outcomes, their behaviour changes. If the evaluation is done participatively the indicators can be quite intangible yet still useful.

Source: Personal communication with Bob Dick December 1997, see also Dick (1991)

12.4. THE THREE RINGS OF EXTENSION EVALUATION

The three rings approach (see Figure 11) is a framework presented in a paper by Jeff Coutts at the 2nd Australasian Pacific Extension Conference in November 1997:

- the inner ring is the project process, activities and outputs
- the middle ring comprises of the communities being directly affected by the extension intervention - with the recognition that other agencies and programs are also having an impact in this ring.
- the outer ring represents the broader community/society where the impact can be measured at a macro level.

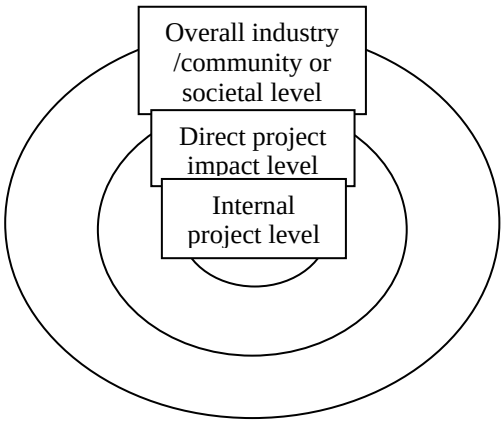


Figure 11 The three rings of extension evaluation

The inner ring – internal project level
According to this framework, the internal project level looks at the internal process and activities of the project over which the project team has some control. This level is similar to Bennett's levels 1-3 and includes activities, participation, inputs and outputs.

The middle ring – the direct project impact level
Coutts (1997) explains that as the second ring is entered the level of complexity and potential for confusion rises significantly. It becomes difficult to attribute the effects of the project intervention directly to the changes occurring, as the changes could be influenced by other programs of events. He suggests that a 'synergy matrix' be developed to highlight the relationship between the extension program and other projects, agencies or groups impacting on the outcomes of the program and to identify the specific knowledge processes which the program itself is contributing towards achieving, amongst the overall desired outcomes. An example of a synergy matrix is shown in Table 9.

The outer ring – overall industry, community or societal level
Coutts explains that most projects cannot lay full claim to observable outcomes at this level. The emphasis here is on measurable changes in tangible benchmarks, for example levels of adoption, levels of sustainable indicators, net farm incomes, and indicators of product quality. He states that due recognition should be given to the role played by other complementary and competing interventions, encouraging and

inhibiting government policies, climatic, market and economic conditions. Describing these provides an overall context for the evaluation of a specific program or project. The role that the program is playing in contributing towards change is clarified, and the levels of changes in benchmarks are put into a realistic timeframe.

Coutts (1997) encourages the evaluator to assess the extent to which the project activities are complementary to, or adding value to, the knowledge processes provided by other parties, and how well these specific knowledge processes are being undertaken. This process can help to identify gaps in the provision of information or other support, and to reduce duplication of effort by different agencies.

Table 9 Example of a simplified synergy matrix for the cotton industry

Knowledge Process	Consultant	Extension	Growers' organisation	Chemical companies
Day –to-day operations, insect monitoring, chemical decisions	XXXXX	X	X	X
Undertaking local trials and validation of new strategies	X	XXXX		X
Demonstrations , workshops , field days and seminar on issues	X	XXXX	X	XX
Number of X's indicate the degree to which the stakeholder was involved in a specifically identified knowledge process ***Source: Summarised from Coutts 1997***				

12.5. INTEGRATED FRAMEWORKS THAT COMBINE SEVERAL FORMS OF EVALUATION

Case Study 16 uses a combination of different evaluation models and frameworks to meet the program's requirements for evaluation. The framework developed was influenced by the model of utilisation focused evaluation and uses the frameworks of Bennett's hierarchy and program logic in understanding the causal links in the program.

Case Study 16
Target 10 framework for project evaluation

Evaluation	*Soft* ◄		●		► *Hard*
Form	*Formative*	●		*Summative*	●
Owen's form	*F5*	*F4* ●	*F3* ●	*F2* ●	*F1* ●
Data collection	*Qualitative*		*Both* ●		*Quantitative*
Stage of program	*Forming*	*Developing* ●		*Settled* ●	*Finished* ●
Program planning level	*Regional*		*State* ●		*National*

The program

The Target 10 extension project was established in 1992 and originally aimed to achieve a substantial improvement in pasture consumption on dairy farms through extension activities. Since then, at the direction of farmers, additional programs have been delivered to address issues besides pasture consumption. The current project goal is 'to secure Victoria's future through a competitive based dairy industry based on efficient and sustainable farm management practices'. Programs, each with specific adoption goal, are modified to accommodate regional differences. For the dairy industry this project represented a new approach in terms of increased participation of stakeholders in the project management and delivery. The Target 10 project combines 4 basic approaches: education, on-farm adoption, community support and benchmarking. This case study relates to the evaluation of the dairy cow nutrition program.

The evaluation

While the evaluation process was being tested certain principles of evaluation evolved including:

- *that evaluation should promote learning and understanding of process and impacts*
- *that it must complement any other evaluation activities taking place within the department, such as routine cost–benefit analysis*
- *that data collection and analysis should only occur when the findings can be and are used*
- *that data analysis not acted upon is a waste of time.*

Goal of Evaluation

The goal of this evaluation was two-fold: it aimed to provide information to help refine the program at the same time as providing evidence that outcomes had been met. The evaluation process was thus designed to drive change rather than to examine whether the project had been successful or not. A utilisation-focused approach was used.

Classification of evaluation

This case study describes a strategic framework for evaluation of a program that uses more than one form of evaluation and includes design clarification (Form 4), process evaluation (Form 3) and impact evaluation (Form 5).

Actors

External consultants were involved in developing the evaluation system, but the process itself is operated from within the program. The audience was largely internal including field staff and managers. However, the summative aspects of the evaluation were also directed at the funders of the project.

Framework for evaluation

A nine step framework (see below) was used to guide the design and implementation of evaluation within Target 10.

1. Audit past dairy cow nutrition extension
(What are the lessons learned from past experience?)

↓

2. Define what success looks like
(Describe the intended observable outcomes of the program)

↕

3. Define the program goal
(Define the main message(s) or extension product)

↓

4. Describe the activities to achieve the goal

↕

5. Define the program logic to identify assumption
Bennett's hierarchy was used to describe the sequence of causes and effects from inputs to final outcomes

↕

6. Design the key evaluations questions
What are the evaluation questions stakeholders want answered?

↓

7. Trial and implement the evaluation methods
Key evaluation questions were formulated on the basis of the program logic as well as discussions with farmers, DRDC, Target 10 operations management and extension officers, and secondly program logic. A variety of evaluation tools were then tested and used to address the key evaluation questions.

↕

8. Describe the roles, responsibilities and resources for evaluation

↓

9. Design the meta-evaluation.
The key questions for the meta-evaluation included:
To what extent and in what ways did the evaluation improve the nutrition program?
What did evaluation users, including those who planned and undertook the evaluation, learn as a result of carrying out the evaluation?

Use of findings
The evaluation findings have been used as the basis to establish a plan of action for delivery over the next 2 years. A second output is a streamlined, more user-friendly evaluation plan with modified methods.

Source: Taken from Bron McDonald 1997

13. SUMMARY GUIDE TO APPROACHES AND METHODS IN EVALUATION

The Review was structured to inform the reader primarily about approaches to evaluation, rather than methods for evaluation. The assumptions behind this thinking are that:

- the aim and purpose of an evaluation should decide the method, rather than the reverse
- there is no shortage of information on both qualitative and quantitative methods in the literature on program evaluation
- a methods-oriented approach to evaluation can be misleading.

Nevertheless, the new evaluator will require guidance in seeking methods to suit particular purposes and situations in evaluation. Tables 5, 6 and 7 are intended to guide the reader to literature on particular methods, models and frameworks used in evaluation, once the aims and functions are clear. These lists are not exhaustive but include reference to the main methods and approaches (and frameworks) that are being used in evaluating agricultural extension in Australia. Many of the methods applied in evaluation come from the applied social sciences, and people new to evaluation are advised to obtain a well-known text on social science research methods, such as Bouma (1996). For evaluation terminology, a useful reference is the *the Thesaurus of Evaluation* by Scriven (1991).

Table 10 Strategic frameworks to conduct evaluations for a program

Terms and brief description of methods/frameworks for evaluation	Source of further information (see Reference list for details)
Bennett's hierarchy: A framework that sets out the hierarchical goals of agricultural extension. It can be a very useful tool for clarifying what aspects of a program are to be evaluated.	Bennett, C.F. (1977). *Analysing impacts of extension programs*. Also mentioned in Van den Ban and Hawkins (1988) *Agricultural Extension*.
CIPP: A systems based framework used in planning evaluation, rather than a method. It is based on the assumption that it is better to improve than to prove. The 4 approaches to evaluation **(context, input, process, product**) bear some resemblance to Owen's 5 forms.	Stufflebeam (1991) *Evaluation Models.*
Logical frameworks are used for both planning projects and monitoring activities. Frequently used in social and natural resource projects in developing countries.	AIDAB (1992) Country Programs Operations Guide/ USAID (1973) *Modifications based on experience.* Farrington and Nelson (1997) Use in PRA.
Owen's five forms of program evaluation: A meta-model of approaches to evaluation useful in planning. The 5 categories of evaluation are (1) Evaluation for program development, (2) evaluation for program design clarification, (3) process evaluation, (4) evaluation in program management and (5) impact evaluation.	Owen (1993) *Program Evaluation, Forms and Approaches.*
Snyder model: A systems-based approach to program evaluation. A main feature is that outputs are divided into 3 time perspectives: (1) immediate effects of project activities (2) targets (present goals) (3) vision or ultimate goal. Used to guide simultaneous outcome and process evaluation and can be participatory or non-participatory.	Dick (1991**)** *Evaluation as action research.*

Table 11 Conceptual models of evaluation

Terms and brief description of evaluation "models"	Source of further information (see Reference list for details)
Action research Pursues both action and research outcomes. It is most commonly done with high levels of participation. Aims to bring about change in such a way that there is a contribution to understanding of the system being changed and the methods used.	Kemmis and McTaggart (1988). *The Action research Planner.* Dick (1991) *Evaluation as Action Research.*
Empowerment evaluation: Use of evaluation concepts, techniques and findings to foster improvement and self-determination. It is designed to help people help themselves and improve their programs using a form of self-evaluation and reflection.	Fetterman, Kaftarian, and Wandersman (1996) *Empowerment Evaluation*: K*nowledge and Tools for Self-Assessment and Accountability*.
Goal -free evaluation: In the pure form of this model of evaluation, the evaluator is not told the purpose of the program but does the evaluation with the purpose of finding out what the program is actually doing.	Scriven (1991). *Evaluation Thesaurus.*
Intervention-oriented evaluation: Designed to support, reinforce and enhance attainment of desired program objectives. Data collection is integrated into program delivery and management. Program staff know the criteria being used to judge success. Feedback is used to increase participant goal attainment as well as program goals. Evaluation is part of the program/intervention model.	Patton (1997) *Utilisation Focused Evaluation.*
Naturalistic inquiry: Evaluation that attempts to arrive at naturalistic generalisations on the part of the audience, is aimed at non-technical audiences like teachers or the public at large, uses ordinary language, is based on informal everyday reasoning, and makes extensive use of arguments which attempt to establish the structure of reality	Guba (1978) *Towards a Methodology for Naturalistic Inquiry in Educational Evaluation.*
Participatory evaluation: Evaluation in which the clients of the program play an active role in directing the evaluation.	Arnstein (1969) *Ladder of citizen participation.*
Social impact assessment attempts to predict the effects that a program might have on various groups or populations. An attempt is made to identify the long-term results or effects (rather than immediate outcomes) and this may be in terms of 'quality of life' indicators.	Meidinger and Schadiberf (1980) *Social impact assessment as evaluation research*.
Utilisation focused evaluation (as opposed to program evaluation in general) is evaluation done for and with specific, intended primary users for specific, intended uses.	Patton (1997) *Utilisation Focused Evaluation.*

Table 12 Methods of evaluation

Method and description	Source of further information (see References for details)
Ex-ante and ex-post evaluation	DOF (1991) *handbook of Cost–Benefit Analysis.*
Case study evaluation: A method at the opposite end of the spectrum to the large-scale survey. It substitutes rich details of a few cases for evaluative conclusions. Can provide valuable insight, especially when used in combination with survey data.	Stake (1995) *The Art of Case Study Research.* Sechrest, *et al.* (1996) *Effective and Persuasive Case Studies.*
Cost–benefit analysis: Used to estimate the $input:$output ratio of a program. Only possible where all costs and benefits can be converted into money terms. Answers the question, "is the $ value of outputs of a program worth its $ cost?" Commonly used in impact evaluation, but can be used in formative studies to assess impact to date (or in ex-ante projections of likely impacts in project proposals).	DOF (1991) *Handbook of Cost–benefit Analysis*. Thompson (1980) *Benefit cost analysis.* Sugden and Williams (1985) *The Principles of Practical Cost–benefit Analysis.* Also – computer packages on cost–benefit analysis.
Cost-effectiveness analysis: Goes beyond cost–benefit analysis - to determine what a program costs against what it does. Used where outputs cannot be reduced to a single dimension (e.g., money). Asks, "are the outputs worth the costs?"	DOF (1991) *Handbook of Cost–benefit Analysis*
Delphi technique: A form of interactive (postal) surveying that utilises iterative questionnaire and feedback approach, and provides participants with an opportunity to revise earlier views based on the response of other participants – until some desired level of consensus is reached. Valuable when participants cannot meet face-to-face. Can produce quantitative data amenable to statistical analysis.	Cary and Salmon (1976) *Delphi and Participatory Planning: focussing the planning process in an agricultural service organisation*.
Experimental evaluation design: Involves pre- and post-test designs. Involves randomly assigning program participants to groups to minimise bias– so that statistical analysis can be conducted. See also: **Quasi-experimental designs.** Used when not possible to randomly select the sample population.	Cambell and Stanley (1963). *Experimental and Quasi-Experimental designs for Research and teaching.* Bouma (1996).
Focused groups gather people together to discuss a topic, usually in the presence of a facilitator to pose questions to 'get the ball rolling' and to help the discussion stay on track.	Krueger (1988) *Focus groups: a practical guide for applied research.*
Goal attainment scaling: Involves the development of outcome scales specifically tailored to the individual, group, agency or system whose program is to be measured.	Kiresuk *et al*. (1994) *Goal Attainment Scaling: applications, theory, and measurement.*
Participant observation: A form of inquiry in which the inquirer (the observer) plays two roles – firstly as an observer and secondly as a genuine participant, having a stake in the group's activities.	Egon and Lincoln (1981) *Effective Evaluation: Improving the Usefulness of Evaluation Results.*
PLA (Participatory Learning and Action): A concept that evolved from PRA. It differs from PRA in that there is a greater emphasis on action – on the basis that it is not enough to simply appraise a rural situation but it is also necessary to facilitate action.	Pretty, Guijit and Scoones (1995) *Participatory Learning and Aciton: a Trainer's Guide.*
PRA (Participatory Rural Appraisal): Uses similar methods to RRA, but the participants become more involved in the inquiry. A main difference is in the behaviour of the researchers. In PRA the researcher aims to facilitate for participants to elaborate their own research and 'hands over the stick'. The contention is that this process can be empowering to the participants.	Pretty, Guijit and Scoones (1995) *Participatory Learning and Aciton: a Trainer's Guide*

Pre-post design: Used when a pre-treatment measure can supply useful information. Many evaluation designs use pre-tests as benchmarks to report the change that has occurred from before the program to its conclusion. The danger is that, unless a control group are considered, there is no means of differentiating between change that occurs in the group naturally and as a result of the program (i.e. through maturation, other learning and intervening events).	Cambell and Stanley (1963) *Experimental and Quasi-Experimental Designs for Research and teaching.*
Qualitative methods in general: Include naturalistic inquiry, open ended questions, participant observation.	Patton (1987) *How to Use Qualitative Methods in Evaluation.*
Questionnaires: Used in formal surveys, or in interviews, or as 'interview guides' to informal surveys. Can be developed to explore or measure attitudes, opinions, behaviour, level of knowledge, life circumstances or other issues.	Many texts on survey & questionnaire design, e.g. Berdie, Anderson and Niebuhr (1986) *Questionnaires; Design and Use.*
RRA (Rapid Rural Appraisal): Generally a 'repertoire of techniques' aimed at collecting information efficiently and involving use of innovative and visual techniques. In Australia, RRA has come to have more specific meaning, i.e. a process of interviewing for an intensive short period (about one week).	Beebe (1985) *Rapid rural appraisal.* Van Beek, P. and Nunn, J. (1995) *Introduction to Rapid Appraisals* SyTREC.
Semi-structured interviews: Used in informal guided interviews where only some questions are predetermined and new questions are developed from discussion. An interview guide is used to ensure that the same format and topics are covered with each respondent.	A user's guide for Participatory Learning and Action *Participatory Methodology Series*. London: International Institute for Environment and Development.

14. REFERENCES

Agricultural Industries Team (1997) "Moving forward." Natural Resources and Environment, Victoria, Melbourne.

AIDAB (1992) Country Programs Operation Guide: Volume 2. Sections 2 and 4 in The Management of Projects, Commonwealth of Australia.

Allen, W.J. (in press) Towards improving the role of evaluation within natural resource management R&D programs: The case for learning by doing. *Canadian Journal of Development Studies*.

Anderson, F.M. and Dillon, J. L. (1985) Farming systems research in the IARCs and other international groups. In J.V. Remenyi (ed.), *Agricultural Systems Research for Developing Countries*, ACIAR Proceedings 11, 141-147.

Appleyard L. (1996a) Target 10 Project Evaluation. In M. Lew (ed.) *Agriculture and Food Initiative: Dairying*. Evaluation Report: Completed Projects 1991-94. Department of Agriculture, Energy and Minerals, Melbourne. 31-40.

Appleyard L. (1996b) APPRAISAL. A Spreadsheet for Evaluating Returns to Agricultural Research. Working Paper 9609. Performance Evaluation Division, Economics Branch, Department of Natural Resources and Environment, Melbourne, Victoria.

Arnstein, S. (1969) Ladder of citizen participation, *American Institute of Planners Journal,* 35, 216-224.

Bardsley, J. B. (1981) *Farmers' Assessment of Information and ITS Sources; An investigation using interactive computer techniques*, PhD Thesis, University of Melbourne, Melbourne.

Bawden, R. J. (1991) Systems thinking and practice in agriculture. *Journal of Dairy Science*, 74, 2362-2373.

Bebbington, A. J., Merrill-Sands, D. and Farrington, J. (1994) "Farmer and Community Organisations in Agricultural Research and Extension: functions, impacts and questions." *ODI Agricultural Administration (Research and Extension) Network paper* (47).

Beebe, J. (1985) Rapid Rural Appraisal; A Critical First Step in a Farming Systems Approach to Research. Networking Paper **5**, Farming Systems Support Project, University of Florida/ USAID, Gainsville.

Bell, J. (1983) *Contemporary Social Welfare*, Macmillan, New York. Cited by Patton, M.Q. (1997) Utilization Focused Evaluation, Sage, Thousand Oaks.

Bennett, C. (1975) Up the Hierarchy. *Journal of Extension* (March / April), 6-12.

Bennett, C.F. (1977) Analysing impacts of extension programs. Washington D.C. Extension service, US Dept of Agriculture.

Berdie D.R., Anderson J.F. and Niebuhr M.A. (1986) *Questionnaires: Design and Use* (2nd edition), Scarecrow Press, London.

Boomsma J. (1996) Target 10: Project Team's Report. In M. Lew (ed.) Agriculture and Food Initiative: Dairying. Evaluation Report: Completed Projects 1991-94. Department of Agriculture, Energy and Minerals, Melbourne. pp. 27-30.

Bouma, B. D. (1996) (3rd edition) The Research Process, Oxford University Press.

Brizius, J.A. and Campbell, M.D. (1991) *Getting Results: A guide for Government Accountability,* Washington DC: Council of Governor's Policy Advisers. Cited by Patton, M.Q. (1997) Utilization Focused Evaluation, Sage, Thousand Oaks.

Cambell, D.T., Stanley, J.C. (1963) *Experimental and Quasi-Experimental Designs for Research and Teaching,* In N.L Gage, *Handbook of Research on Teaching,* Rand McNally, Chicago.

Cary, J. W. (1993) Changing Foundations for Government Support of Agricultural Extension in Economically Developed Countries, Sociologia *Ruralis*, 18, 336-347.

Cary, J.W., Salmon, P.W. (1976) *Delphi and Participatory Planning: focusing the planning process in an agricultural service organisation*. Agricultural Extension Research Unit, School of Agriculture & Forestry, University of Melbourne.

Casley, D.J. and Kumar, K. (1987) Project Monitoring and Evaluation in Agriculture. Maryland: The John Hopkins University Press.

Chambers, R. (1983) Rural Development: putting the last first. Harlow: Longmann.

Chambers, R. (1994) The origins and practice of participatory rural appraisal, World Development, 22, 953-969.

Chen, H.T. (1990) *Theory-Driven Evaluation*. Newbury Park: Sage Publications.

Clark, R.A., Bourne, G.F., Cheffins, R.C., Esdale, C.E., Filet, P.G., Gillespie, R.C. & Graham, T.W.G. (1996) The Sustainable Beef Production Systems Project: Beyond Awareness to Continuous Improvement. Queensland Department of Primary Industries, Brisbane. Project Report Series Q96002.

Conway, G. (1985) Agricultural ecology and farming systems research, In J. V. Remenyi (ed.) *Agricultural Systems Research for Developing Countries*, ACIAR Proceedings 11, Canberra, 43-59.

Coutts, J. (1997) Exposing the tiger – a fresh look at the evaluation of extension. Paper read at 2nd Australasian Pacific Extension Conference, at Albury, New South Wales.

Cronbach, L.J. (1963) Evaluation for course improvement. In *New curricula*, edited by R. Heath. New York: Harper & Row.

Dart, J.J. (1997) Facilitating tribal farmers to identify change to their farming systems using participatory farming systems models. Paper read at Australasian Evaluation Society International Conference, Adelaide, South Australia.

Davis, D. (1994) An evolutionary approach to facilitating organisational learning: An experiment by the Christian Commission for Development in Bangladesh. A WWW publication, Swansea: Centre for Development Studies. http:/www.swan.ac.uk/cds/rdslp.htm.

Department of Agriculture (1993) Corporate Plan. Melbourne: Department of Agriculture Victoria.

Dick, R. (1991) *Evaluation as Action Research*, Interchange (Resource Document), 37 Burbong St, Chapel Hill, Australia.

Dillon, J.L. (1965) Farm management in Australia as an academic discipline, Review of Marketing and Agricultural Economics, 33, 175-189.

DOF (1991) *Handbook of Cost–benefit Analysis*, Department of Finance, Australian Government Publishing Service, Canberra.

Egon G.G. and Lincoln, Y.S. (1981) Effective Evaluation: Improving the Usefulness of Evaluation Results through Responsive and Naturalistic Approaches. Jossey-Bass, San Francisco

Farrington, J. and Nelson, J. (1997) "Using Logframes to Monitor and Review Farmer Participatory Research". Network Paper No. 73, Overseas Development Institute, London.

Fetterman, D.M., Kaftarian, S.J. and Wandersman, A. (1996) Empowerment Evaluation: Knowledge and Tools for Self-Assessment and Accountability, Sage, Thousand Oaks, California.

Fishpool, K. I. (1993) Quality Assurance Evaluations. An Outline of Principles/Best Practice for Program Evaluators in NSW Agriculture, NSW Agriculture.

Foale, M.A. (1997) Management of dryland farming systems in North Eastern Australia. How can scientists impact on farmer's decision making? *Agricultural Science* **10**, 34-37.

Freeman, H. E. (1977) The Present Status of Evaluation Research, *Evaluation Studies Review Annual*, M.Guttentag and S.Saar (eds), Sage, Beverly Hills.

Friere, P. (1972) Pedagogy of the Oppressed, Harmondsworth: Penguin.

Geertz, C. (1973) The Interpretation of Cultures. USA: Basic Books.

Gredler, M. E. (1996) *Program Evaluation*, Prentice-Hall, inc., Englewood Cliffs, New Jersey.

Guba, E. G. (1978) Towards a methodology of naturalistic inquiry in educational evaluation. CSE Monograph Series in evaluation No. 8, Centre for Study of Evaluation, Los Angeles.

Guba, E. G. and Lincoln, Y. S. (1981) *Effective Evaluation*, Jossey-Bass, San Francisco.

Guba, E. G. and Lincoln, Y. S. (1989) *Fourth Generation Evaluation*, Newbury Park: Sage.

Hall, H. (1994) Reducing the risk – integrated pest management in potatoes. Proceedings of National Crisping Potato Industry Workshop, Mildura, July 1995. pp 85-90.

Hart, E. Bond., M. (1995) Action Research for Health and Social Care: A Guide to Practice, Open University Press, Buckingham.

Holter, I.M, and Schwartz-Barcott, D. (1993) Action Research: What is it? How has it been used and how can it be used in nursing? *Journal of Advanced Nursing* 18 (18), 289-304.

Horne, P. A. (1990) An integrated pest management strategy for potato moth. Proceedings of the 5th national Potato Conference, June, Warragul, Victoria, pp. 60-63.

House, E.R. (1978) Assumptions underlying evaluation models. *Ed. Researcher* 7 (3), 4-12.

House, E.R. (1980) *Evaluating with Validity*. Beverly Hills: Sage.

Hulme, D. (1997) Impact Assessment Methodologies for Microfinance: A review. Manchester.

Jones, S., Khare, J.N., Mosse, D., Sodhi, P.S., Smith, P., Witcombe, J.R. (1996) The KRIBHCO Rainfed Farming Project: An approach to participatory farming systems development. Swansea: Centre for Development Studies, University of Wales.

Kallenberg, R. (1990) The Construct Turn in Sociology, Cited by Holter *et al.* (1993). Oslo: Oslo Institute for Social Research.

Kalpan, A. (1994) The conduct of Inquiry, Chandler, San Francisco.

Kemmis, Stephen and Robin Mctaggart (1988) The Action research Planner, Deakin University, Geelong.

Kiresuk, T.J., Smith, A., Cardillo, J.E. eds. (1994) *Goal Attainment Scaling: applications, theory, and measurement*. Hillsdale, N.J.: Erlbaum Associates.

Krueger, R.A. (1988) Focus Group Interviews: A practical guide for applied research. Newbury Park: Sage.

Kolb, D.A. (1984) Experiential Learning: Experience as a Source of Learning and Development. Englewood Cliffs, NJ, USA.

Lazarsfeld, P. F., Rosenburg, M. (1955) The Language of Social Research. Cited by Shadish *et al.* (1995), Free Press, Glencoe, IL.

Lees, J. W. (1991) More than accountability. The Rural Development Centre, University of New England, Armidale.

Lew, M. (1996) Agriculture and Food Initiative: Dairying. Evaluation Report: Completed Projects 1991-94. Department of Agriculture, Energy and Minerals, Melbourne.

Lewin, K. (1948) Resolving Social conflicts: Selected papers on group dynamics, Cited by Shadish *et al.* (1995), Harper and Brothers, New York.

Madaus, G. F., Scriven, M., and Stufflebeam, D. L. (1991) Evaluation Models, Kluwe-Nijhoff Publishing.

Macadam, R.D. (1996) Developments in agricultural industries research and development in Australia with particular reference on extension and the relevance of systematic developments as a response to the current situation. Country paper presented at Asia Pacific Seminar on Education for Rural Development, Obihiro, Japan.

McCown, R.L. (1991) Research in a farming systems framework. In V.R. Squires and P.G. Tow (eds), Dryland Farming: A Systems Approach. An Analysis of Dryland Agriculture in Australia, Oxford University Press, Melbourne. pp. 242-249.

Meidinger, E ., and Schadiberg, A. (1980) Social Impact Assessment as Evaluation Research. *Evaluation Review* 4 (4), 507-535.

Morrison, D. (1993) Square pegs in round holes: perspectives of cost–benefit gained from evaluation of Western Australia Department of Agriculture activities. In Proceedings of a workshop held in Canberra on 25-26 May 1993.Gains for Grains, 4. Canberra, 35.

Mosse, D., Ekande, T., Sodhi, P.S., Jones, S., Mehta, M., Moitra, U. (1995) Approaches to Participatory Planning: a Review of the KRIBP Experience. Swansea: Centre for Development Studies, University of Wales.

Owen, J. M. (1993) Program Evaluation, Forms and Approaches, Allen and Unwin, St Leonards, NSW, Australia.

Patton, M. Q. (1987) How to Use Qualitative Methods in Evaluation, Sage, Newbury Park, USA.

Patton, M. Q. (1997) Utilization Focused Evaluation, Sage, Thousand Oaks, USA.

Pretty, N.P., Guijit, I. and Scoones, I. (1995) Participatory Learning and Action: a Trainer's Guide, IIED Participatory Methodology Series. London: International Institute for Environment and Development.

Purcell, D (1984) Agricultural Extension: Lessons from Completed Projects, Draft Report no 13000, Operations Evaluation Dept, World Bank, Washington.

Rhoades, R. and Boothe, R. (1982) Farmer-back to farmers: A model for generating agricultural technology, *Agricultural Administration*, 11, 127-137.

Riches, J. R. H. (1973) Evaluation of a Farm Business Management Extension Program, Glenormiston Agricultural College, Noorat, Victoria.

Rossi, P. H., Freeman, H. E., Write, S. R. (1979) Evaluation: A systematic Approach, Sage publications, Beverly Hills, London.

Russell, D.B., Ison, R.L., Gamble, D.R. and Williams, R.K. (1989) A critical review of rural extension theory and practice, Australian Wool Corporation, Sydney.

Scriven, M. (1967) The methodology of evaluation. In *Curriculum evaluation.* Edited by Stake. Chicago: Rand McNally.

Scriven, M. (1991) Evaluation Thesaurus (4th Edition) Newbury Park: Sage.

Sechrest, L., Stewart, M., Stickle, T.R., and Sidani, S. (1996) Effective and Persuasive Case Studies. Toolkit commissioned by the Evaluation Centre at Human Services Research Institute, Cambridge, MA.

Shadish, W. R., Cook, T. D. and Leviton, L. C. (1995) Foundations of Program Evaluation, Sage, Newbury Park, California.

Smith, M.F. (1989) Evaluability assessment : a practical approach, Evaluation in education and human services. Boston: Kluwer Academic.

Smith, N.L. (1994) Evaluation Models and Approaches. In *International encyclopedia of Education*, edited by T. Husen and Postlethwaite. Oxford: Elsevier Science.

Stake, R.E. (1967) The Countenance of Educational Evaluation. Teachers College Record 68, 523-40.

Stake, R.E. (1973) Program evaluation: particularly responsive evaluation. In *Evaluation Models*, edited by G. F. Madaus, M. S. Scriven and D. L. Stufflebeam. Boston: Kluwer-Nijhoff.

Stake, R.E. (1995) The Art of Case Study Research, Sage, Newbury Park, California.

Stewart, L. (1994). A review of the monitoring and evaluation of property management planning in South-East Queensland: Department of Primary Industries.

Straw, W. (1997) Personal communication. Department of Natural Resources and Environment, Attwood, Victoria.

Straw, W., Winston, J. and Boomsma, J. (1997). An extension approach to developing skills in project evaluation. Paper read at 2nd Australasian Pacific Extension Conference, at Albury, New South Wales, Australia.

Stufflebeam, D. (1994) Empowerment Evaluation, Objectivist Evaluation, and Evaluation Standards: where the future of evaluation should not go and where it needs to go. *Evaluation practice*, 15(3), 321-38.

Stufflebeam, D. (1991) Evaluation Models. In Madaus, G. F., Scriven, M., and Stufflebeam, D. L. (1991), Kluwe-Nijhoff Publishing.

Stufflebeam, D. and Webster, W.J. (1981) An analysis of alternative approaches to evaluation. In *Evaluation Studies Review Annual*, edited by H. E. Freeman and M. A. Solomon. Beverly Hills: Sage.

Sugden, R. and Williams, A. (1985) the Principles of Practical Cost–benefit Analysis, Oxford University Press, Oxford, UK.

Thompson, M.S. (1980) *Benefit Cost Analysis for Program Evaluation*, Sage, Beverley Hills, California.

Tyler, R. W. (1967) Changing the Concepts of Educational Evaluation. *Perspectives of Curriculum Evaluation*, R. E. Stake (ed.), Rand McNally, New York.

USAID (1973) The Logical Framework: Modifications Based on Experience, United States Agency for International Development, Washington D.C.

Van Beek, P. and Nunn. J. (1995) *Introduction to Rapid Appraisals,* SyTREC Pty Ltd, 9 Carlock Promenade, Karalee Queensland 4306

Van den Ban, A. W. and Hawkins, H. S. (1996) (2nd Edition) *Agricultural Extension*, Blackwell Science, Berlin, Germany.

Wadsworth, Y. (1997) *Everyday evaluation on the run*. Melbourne: Action Research Issues Association (incorporated).

Watson, A., Hely, R., O'Keefe, M., Cary, J., Clark, N. and Allen, J. (1992) Review of Field-based Services in the Victorian Department of Food and Agriculture, Melbourne, Victoria.

Weiss, C.H., and Bucuvalas, M.J. (1977). The challenge of social research to decision making. In *Using Social research in public policy making*, edited by C. H. Weiss. New York: Lexington.

Wholey, J.S. (1983) *Evaluation and Effective Public Management*. Little Brown: Boston.

Williams, D. B. (1968) *Agricultural Extension*. Melbourne University Press, Melbourne.

Winston, J. (1997). Personal communication. Royal Melbourne Institute of Technology, Bundoora Campus, Melbourne.

Winston, J.A. (1991). "Linking Evaluation and Performance Management", review paper. Paper read at National Evaluation Conference, at Adelaide.

Woods, E., Moll, G., Coutts, J. Clark, R. and Irvine, C. (1993) Information Exchange: A Review. LWRRDC, Canberra.

Worthen, B. R., Sanders, J. R., and Fitzpatrick, J. L. (1997) Program Evaluation: Alternative Approaches and Practical Guidelines, Longman, New York.

15. GLOSSARY

Accreditation – the determination of the adequacy of the developed plans and procedures of an organisation for the future delivery of services (House 1980).

Benchmarking – identification of the practice and performance of like organisations known to be leaders in a given field, the development and monitoring of actions designed to implement and monitor the adopted practice (Camp, 1989).

Formative evaluation – formative evaluation has the purpose of improving programs.

Model – (in evaluation) has been used with wide variability and considerable ambiguity in evaluation practice and literature. Generally it refers to particular conceptions, approaches, methods and even loose theories for thinking about and/or conducting evaluations (Smith 1994).

Monitoring – is a part of evaluative inquiry; it is a data collection process that continues throughout the life span of the program, and is generally a tool of program management. Monitoring uses indicators to monitor change over time. (However, many organisations have their own definitions of this term).

Naturalistic inquiry – Scriven (1991) defines naturalistic evaluation (or methodology) as an approach that minimises much of the paraphernalia of science, for example; technical jargon, statistical inference and the effort to separate the observer from the subject and experimental design. Instead there is a focus on the use of metaphor, analogy, informal (but valid) inference, vividness of description, reasons-explanations, multiple perspectives and tacit knowledge. Naturalistic inquiry employs qualitative methods. Advocates include Stake (1967) and Guba (1978).

Performance indicators – are criteria to be monitored over time to give an indication of performance of some aspect of the program.

Process evaluation – involves the collection of information about program activities.

Process-oriented program – a program that does not have specifically defined goals at the onset.

Program – is a set of planned activities directed toward bringing about a specified change' in an identified and identifiable audience' Smith 1988.

Program evaluation – the process of providing information designed to assist decision making about the object being evaluated. The term 'evaluand' has been used as an alternative generic term to 'object'. In practice the 'evaluand' could be a program, a policy, an event, or a training manual or text book' (Owen 1993).

Project – are time-bound efforts and often within a program (Scriven 91).

Quantitative data analysis – in evaluation usually refers to an approach involving a heavy use of numerical measurement and data analysis methodology from social science or from accounting. It focuses on the testing of specific hypotheses that are smaller parts of some larger theoretical perspective (Scriven 1991).

Qualitative data analysis – in evaluation refers to the part of evaluation that can't be usefully reduced to quantitative measures. A large part of good evaluation is wholly or chiefly qualitative, meaning that description and interpretation make up most of it (Scriven 1991).

Summative evaluation – is a purpose distinction and has the purpose of informing overall judgments of merit or worth so that summative decisions can be made.

MAIN ACRONYMS USED

CEM	Community environmental monitoring
IKS	Information knowledge system
NGO	Non government organisation
PRA	Participatory rural appraisal
RRA	Rapid rural appraisal
SMART	Specific measurable, accountable, realistic and timebound
WCM	Whole catchment management

APPENDIX

APPENDIX 1. PROGRAM PARTICIPATION INDEX

Score	Description of the level of participation in the program
0	Designed and executed from above, no consultation
2	Designed and construed from above, consultation with clients and modification
4	Designed from above, workshops and large effort to incorporate beneficiary perception
6	Initiated from above, run by community who set their own objectives, facilitated from above
8	Designed in collaboration with beneficiaries, implemented by beneficiaries
10	Designed by beneficiaries for beneficiaries, no outside direction

APPENDIX 2. EVALUATION PARTICIPATION INDEX

Score	Description of the level of participation in the evaluation
0	Evaluated externally, no consultation of evaluation indicators by staff or beneficiaries
1	Externally evaluated, staff consulted and collaborate in development of indicators, beneficiaries not consulted
2	Evaluation carried out by staff, beneficiaries not consulted in design
3	Externally evaluated, staff and beneficiaries consulted in developing indicators
4	Evaluation carried out by staff, beneficiaries consulted in the development of indicators
5	Staff and beneficiaries develop evaluation together
6	Beneficiaries are facilitated to carry out a self-evaluation of the program
7	Beneficiaries evaluated program unfacilitated.

APPENDIX 3. LIST OF CASE STUDIES AND THEIR CONTRIBUTORS

Grateful acknowledgment is made to all those who contributed written or verbal reports to the Review.

Author (s)	Title	Form	no	pp
Allan Curtis and Terry De Lacy (1994)	*Landcare: does it make a difference?*	Impact assessment (Form 5), goal-based approach	1	35
Jim Woodhill, Paul Trevethan, Judy Frankenburg, Celine Beaulieu and the West Hume Landcare group	*A Case Study Undertaken with the West Hume Landcare Group as Part of an International Research Project on Participatory Watershed Development*	Impact assessment (Form 5), needs-based approach	2	37
Boomsma (1996) Appleyard (1996) Appleyard (1996a)	*Cost Benefit Analysis of the Target 10 Dairy Extension Program*	Impact assessment (Form 5), economic analysis of outcomes	3	41
Jessica Dart and the KRIBP team	*Evaluation of a participatory process project in Western India*	Impact assessment (Form 5), evaluation for illumination	4	44
Paul Horne (1997) IPM Technologies P/L, Victoria.	*IPM program for potato pests: an example of monitoring & evaluation*	IPM Technologies P/L, Victoria. Monitoring and evaluation for management (Form 2)	5	51
Bob McCown, CSIRO (Australian Production Systems Research Unit, Toowoomba) and Jeff Coutts (Rural Extension Centre, University of Queensland, Gatton)	*FARMSCAPE (Farmer-Adviser-Researcher Monitoring, Simulation, Communication and Performance Evaluation) for improved management of dryland cropping*	Monitoring, (Form 2 4), aimed at improvement	6	55
Rick Davies, Centre for Development Studies, Swansea, UK	*An evolutionary approach to facilitating organisational learning: An experiment by the Christian Commission for Development in Bangladesh*	Monitoring (Form 4) monitoring without using indicators	7	57
Drafted by John Petheram Ffrom information received from Andrew Almond (Longerenong College) and Nagi Mina (*Farm Pride Foods, Keysborough)*	Quality Assurance for Farm Pride Foods Eggs, Victoria	*Monitoring and evaluation (Form 2)Longerenong College (University of Melbourne) and Nagi Mina, Farm Pride Eggs, Keysborough*	8	61
James Butterworth and Geoff Syme, CSIRO Land and Water, Perth.	*National Whole Catchment Management: Review and Analysis of Processes*	Process evaluation (Form 3), study of implementation at a national level	9	64
Cam Nicholson, DNRE	*Conrongmite Salinity Program and Global Assessment Scaling*	Process evaluation, (Form 3), aimed at monitoring a group progress	10	66
Alison Spencer, Rural Extension Centre, U. Queensland, Gatton	*Improving the Usefulness of the Pigpulse Report, an Evaluation*	Process evaluation, (Form 3), guidance for refinement	11	70

David Beckinsale (1997) Department of Natural Resources and Environment, Melbourne	*A Workshop to Refine the Program Logic for Corangamite Salinity Program*	Design clarification (Form 4), program logic to clarify the design	12	73
N.A. Hamilton (DPI Toowoomba), Blacket and Gus Hamilton from the DPI , Goodndiwindi Queensland (1992)	*Understanding farmer decision making on land use*	Evaluation for program development (Form 5) A needs analysis using participatory approaches	13	77
John Petheram, Longerenong College, University of Melbourne	*A Logframe for Pproposed project to S support Landcare Ggroup A activities in the Wimmera*	Framework for planning evaluation. Or evaluation for project development (Form 5) and for M&E (Form 2)	14	84
Personal communication with Bob Dick, December 1997, see also Dick (1991)	A program to replace imprisonment for some offenders using the Snyder model	Framework for planning evaluation	15	86
Bron McDonald, DRNE	*Target 10*	Framework for planning evaluation	15	91

APPENDIX 4. LIST OF CONTRIBUTORS

Appreciation is expressed to the following people, who contributed materials for use in the Review of evaluation in agricultural extension.

Contributors	**Organisation**
Almond, Andrew	Longerenong College, Horsham
Anderson, Jock	World Bank, Washington
Baxendell, Sandra	QDPI, Toowoomba Qld
Beckingsale, David	DNRE, Melbourne
Bilston, Larissa	Qld DPR with ACAIR
Boomsma, John	DNRE, Ellenbank
Burnside, Don	Agriculture WA, Perth
Butterworth, James	CSIRO, WA
Cameron, D.	UQG
Carr, Anna	ANU, Canberra
Cary, John	Melbourne University
Conelly, D.	DRDC, Dairy Research and Development Corporation, Victoria
Coutts, Jeff	Rural Extension Centre, Gatton, U.Q.
Dare, Jill	Agriculture WA & NLP
Dart, Jessica	University of Melbourne, Parkville
Davis, Rick	Centre for Development Studies, Swansea, UK
Dick, Bob	Southern Cross University & Brisbane
DPIF Alice springs	DPIF, Northern Territory
Dunn, Tony	CSU, Dept. of Agriculture, Wagga Wagga
Falvey, Lindsay	Melbourne University
Fishpool, Ken	NSW Agriculture and Fisheries
Frank, Bruce	UQ, Department of Agriculture, Brisbane

Goldsworthy, David	Economic Evaluation Unit, DNRE, Department of Natural Resources and Environment, Melbourne.
Hamilton, Gus	QDPI, DalbyQueensland
Harris, Ron	DNRE, Bendigo
Herbert, Alan	Agriculture WA, South Perth
Horne, Paul	IPM Technologies, Melbourne
Howard, Kirsty	DNRE, Rutherglen
Hunt, Keryn	QDPI
Hunt, Natalie	NLP
Kington, Liz Kilpatrick, Sue Lawrence, David	University of Western Australia Centre for Research and Learning in Regional Australia, University of Tasmania QDPI, Pittsworth, Qld
Leach. Greg	QDPI Longreach
Lees, Jim	Rural Development. Centre, UNE
Letts, Malcolm	QDPI, Brisbane
Macadam, Bob	University of W. Sydney, Hawkesbury
Marsh, Sally (and Dave. Pannell)	University of W.A
McCown, Bob	APSRU, Toowoomba, Queensland
McDonald, Bron	DNRE, Victoria
Mckinlay, John	Agriculture Victoria
Middleton, Alison	QDPI
Mina, Nagi	Farm Pride FoodsEggs, Keysborough
Nicholson, Cam	DRNE, Geelong
O'Loughlin, J	DPIF
Owen, John	CPE, University of Melbourne
Petheram, John	Longerenong College, University of Melbourne
Simpson, Ian Spencer, Alison	Orange NSW QDPI Moorooka QLD
Stanley, Jane	FOCUS, Mount Crosby Queensland
Stantial, John	Massey University, New Zealand
Stewart, Leath	QDPI, Bundaerg QLD
Stoneham, Gary	DNRE, Victoria
Straw, Warren	DNRE, Attwood
Symes, Geoff Van Beek, Peter Vanclay, Frank	CSIRO, Florest Park, WA Sytrec Pty Ltd. Fax (03) 93449814 Rural Soc., CSU, Wagga
White, Kerrie	QDPI
Wightman, Jane	QDPI
Willy, Allen	New Zealand

PRINTER

Back cover blurb:

Greater emphasis on accountability in the delivery of extension programs has put new focus on techniques for evaluating such programs.

This report discusses the discipline of program evaluation and then reviews five main forms of evaluation and illustrates them with case studies drawn from agricultural extension in Australia.

It is hoped that improved access to sound information on evaluation methods will be reflected in better planning of agricultural extension projects, more effective monitoring and evaluation of programs, and hence more efficient use of industry funds and greater benefits for clients of all research and development programs in rural Australia.

This report is a new addition to RIRDC's range of over 250 research publications and forms part of its Human Capital, Communications and Information Systems R&D program.

Spine text:

Review of Evaluation in Agricultural Extension RIRDC Publication no. 98/136