Strategic Approaches for Evaluation in Australian Cotton Research Programs

A research dissertation submitted in part fulfilment of the requirements for the degree of Masters in Rural Systems Management

by

Ingrid H. Christiansen

Student Number: 33053442

Centre for Rural and Regional Innovation – Queensland

School of Natural and Rural Systems Management

The University of Queensland

June 2004

Statement of original authorship

I hereby state that the work contained in this assignment has not previously been submitted for assessment, either in whole or in part, by either myself or any other student at either the University of Queensland or at any other tertiary institution except where explicitly acknowledged. To the best of my knowledge and belief, the assignment contains no material which has been previously published or written by another person except where due reference is made. I make this Statement in full knowledge of and understanding that, should it be found to be false, I will be subject to disciplinary action under Statute No. 13 and the Examination Rules of The University of Queensland.

Ingrid H. Christiansen

ACKNOWLEDGEMENTS

Particular thanks are due to the cotton industry researchers who gave their time for interviews: Michael Bange, Greg Constable, Angus Crossan, Alice del Socorro, Sandra Deutscher, Martin Dillon, Peter Gregg, Viliama Heimoana, Nilantha Hulugalle, Joe Kochman, Peter McGee, Robert Mensah, Melina Miles, Dave Murray, David Nehl, James Neilsen, Inakwu Odeh, Nick Reid, Grant Roberts, Brad Scholz, John Stanley, John Triantifilis, Willem Vervoort and Lewis Wilson.

Thanks also to those who contributed to the development of this research project, including Dallas Gibb, Bruce Pyke and Guy Roth.

Research guidance, supervision and insights from Professor Beth Woods were greatly appreciated.

Thanks to the Cotton Research and Development Corporation and the Australian Cotton Cooperative Research Centre who funded this research. Thanks also to these organisations and to the Queensland Department of Primary Industries and Fisheries for the flexibility to link this university study with my work program.

ABSTRACT

Research and development programs are under increasing pressure to demonstrate and enhance their impacts towards the triple bottom lines of environment, economic and social criteria. Cotton research, development and extension funders are looking for strategies both to meet this reporting need and to build capability and impact-thinking amongst research and extension providers. This research reviews literature and existing reports and, through a series of unstructured and semi-structured interviews, explores the perspectives of both research funders and researchers towards evaluation.

It has identified a low level of formal evaluation practice and understanding amongst cotton researchers. However, many researchers regularly gather feedback from industry about their research and are willing to further explore constructive evaluation. Strong views are held about a lack of specialised skills, and the need to engage these rather than build them solely within existing staff. Building an appreciation and understanding of evaluation amongst researchers was regarded important to aid evaluation and improve projects.

A strategic, holistic view is needed for evaluation of Cotton RD&E to be efficient and to minimise the pressure on industry in gathering data. This efficiency has three core elements: 1) Finding a balance between projects and programs – it is suggested to look at individual projects up to the level of outputs and at project clusters, programs or key questions for evaluation of outcomes and impacts; 2) minimising the pressure on industry by gathering data in smart & efficient ways, unobtrusively where possible; and 3) Develop and resource a clear evaluation strategy.

Also identified in this research are diverse values and roles for cotton research, with a gradient of embeddedness in industry. There are some conflicts between perceived industry needs and organisational needs for some scientists, particularly about the need for peer reviewed publishing.

Contents

TABLE OF FIGURES

BACKGROUND

Growing demand for accountability of research and development corporations, tied with a requirement to ensure that research influences practice, has led to a demand for evaluation of research and extension programs. Australia's rural research and development corporations (RDC) have identified the need for evaluation of their research portfolios and are in the process of developing mechanisms by which this may happen.

The RDCs are accountable to the Commonwealth through annual reports that justify the expenditure on research and development by detailing outcomes and impacts. Increasingly, this is required to focus on the triple bottom line impacts of economics, environment and social. Additionally, some organisations have legislative requirements to report on elements of economic, environmental and social impacts.

The Cotton Research and Development Corporation (CRDC) and the Australian Cotton Cooperative Research Centre (Cotton CRC) have been working towards measuring impacts of research investments in the cotton industry. Roth (2003) identified a set of measurable sustainability indicators to monitor and report on the TBL impact of the CRDC's research investments. Measurement of these indicators will aid in determining the changes that have occurred in these factors and some of this may be implied to be linked to research. However, such sustainability indicators do not necessarily measure the impact that research programs have had on this change. Understanding the contribution of research towards change in economic, social and environmental factors requires knowledge of effects at multiple levels. Like many other research groups, CRDC and the Cotton CRC now need to develop systems and capability to gather and comprehend this knowledge about the outputs, outcomes and impacts of their research investments.

To date, the development of evaluation capability and procedures in Cotton RD&E has been a somewhat ad hoc response to increasing need. CRDC have taken a lead in their recent Five Year Strategic Plan (CRDC 2003b) and Annual Report (CRDC 2003a) and have also gradually increased exposure of researchers to the concepts of the triple bottom line and evaluation of impacts. Evaluation has been developed further in the National Cotton Extension Network through training, workplans, evaluation studies of outputs and a few of outcomes, the introduction of frameworks and reinforcement from the CRDC Board of the need for evaluation (Christiansen, Pyke *et al.* 2003).

Developing any new systems and approaches will benefit from engaging key stakeholders. Researchers, key stakeholders themselves in any changes to research approaches and reporting, have not yet been actively engaged in developing evaluation process. The main involvement of researchers to date has been through their routine bi-annual reporting to CRDC and Cotton CRC.

A strategic approach will aid in building evaluation processes and capability across Cotton RD&E programs. Such an approach would do well to learn from evaluation experiences in other industries and disciplines and engage stakeholders from the early stages. For this purpose, this study has reviewed experiences in evaluation approaches, capability building and frameworks and the barriers and triggers to their success. It has engaged researchers through interviews to explores their views, concerns, skills and suggestions for TBL evaluation.

An overview of issues and opportunities for building evaluation capability and process is presented by drawing on the literature and the perspectives of researchers and research funders.

Research Questions

The focus for this research was developed through an iterative process together with a number of key stakeholders and participants. This process initially scoped what prior work had been done and what research questions emerged from this. Iterative discussions with informed persons and key stakeholders further developed the key research questions.

The key research questions emerging from this were:

1. ***How do cotton researchers currently approach, think of and practice evaluation of the triple bottom line impacts of their research?***
   What are their research approaches and values that may influence this?

2. ***What strategies may be used to move the triple bottom line evaluation approach forward?***

Research Approach

This research has been approached from the philosophical paradigm of *contructivism*. The ontology that underpins contructivism is that there are multiple realities that exist and that the nature of these depends on the individual. A constructivist approach aims to better understand these multiple realities rather than to identify a particular 'truth' about what reality is (REC 2003).

This approach was chosen as the research investigation aims to better understand the various different perspectives about evaluation and triple bottom line and to look for all opportunities to progress this need. The study does not aim to specifically identify the 'true' or 'best' perspectives, which would be a more positivist approach. The constructivist research approach appreciates multiple perspectives. This approach is considered suitable for this study as it has explored the multiple perspectives of just one sector of stakeholders – the researchers. Within this sector alone exists high diversity, and no doubt other stakeholder sectors will hold further perspectives. The constructivist approach acknowledges that each of these perspectives is valid and that no one position is more 'true' than another.

Additionally, my epistemology as a researcher aligns strongly with the constructivist approach of being involved with the research and in understanding the research questions through experiences. Having been involved with cotton research and extension for three years, and knowing most of the players well, it would be difficult for me to distance myself from this approach.

The main body of theory that has been drawn upon is that of evaluation. To a lesser degree, it links with theories from business, management, research management and adult learning.

Methodologies used in this study are largely qualitative. The validity and rigour of the research has been built by using triangulation, and an iterative process of research development. Potential bias in the research is acknowledged in the discussion of limitations. Methods used in the study included a brief review of existing data, interviews of researchers and a review of the literature.

To understand the 'fit' of this research, it can be considered in terms of Checkland's (1984) FMA framework. As detailed in Figure 1, this research focuses primarily on the area of developing strategic methodologies for evaluation in cotton research programs.

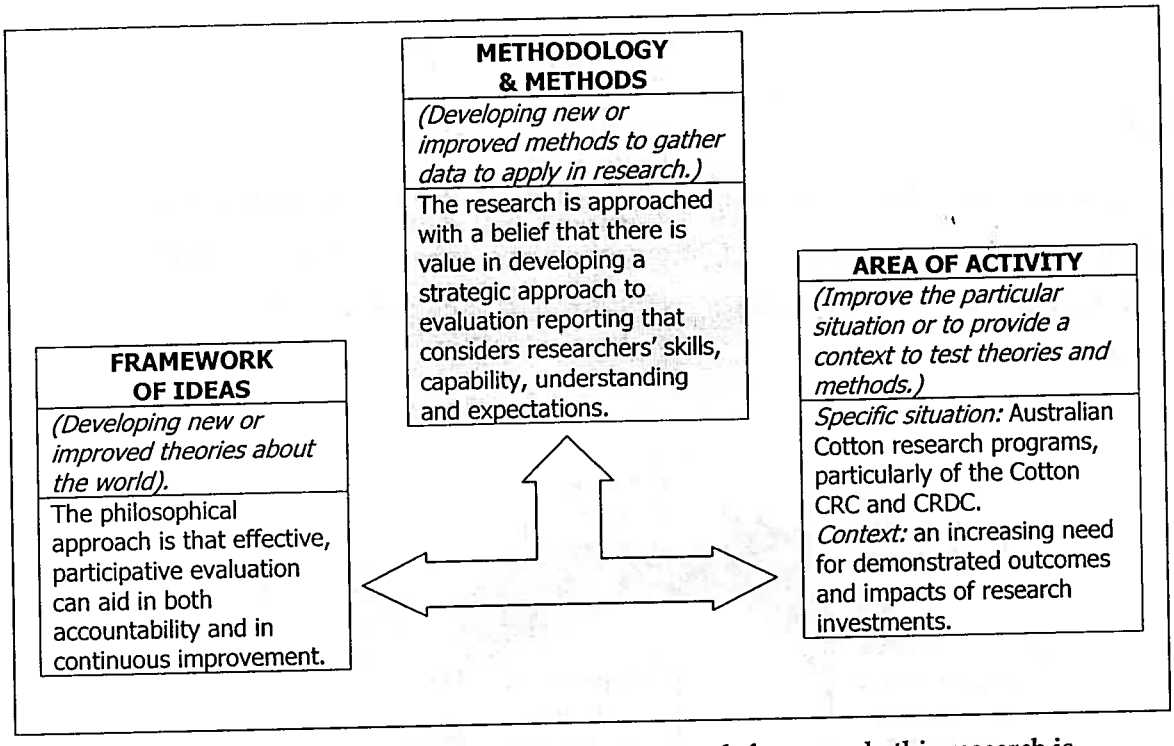


Figure 1 **Positioned in Checkland's (1984) research framework, this research is focussed primarily in the area of developing methods and methodology.**

Outline

This report presents the background to the study, a review of the literature, the methods used and the findings of each element of the research. It subsequently draws on each of these in the discussion of the key research questions and identifies potential areas for future research.

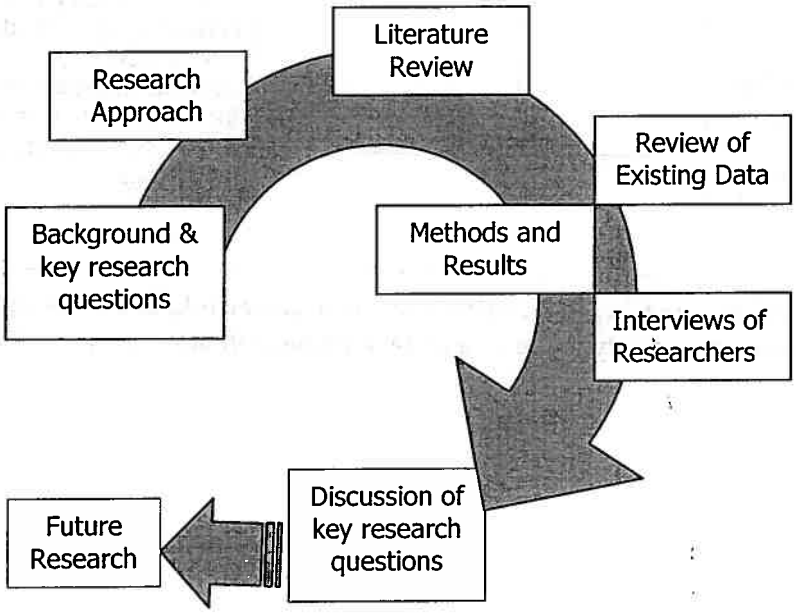


Figure 2 Outline of research report

LITERATURE REVIEW

"Excellence must be combined with relevance in a successful modern public sector research environment". (Batterham 2000)

A global shift towards "knowledge based economies" is recognising the generation and application of knowledge as a key to prosperity. Growing importance is placed on the intangible capital related to knowledge such as research and development, human, structural and social capital and the ability to innovate. The value of knowledge is not entirely new to Australia; research and development, education and other knowledge services have contributed much of the success of Australia's traditional economic base such as primary industries (Batterham 2000).

In a key discussion paper, 'The Chance to Change' (Batterham 2000), Australia's Chief Scientist identified that innovation is essential in order for Australia to participate and thrive in this knowledge economy. Batterham advocates 'investment' rather than 'expenditure' in science, engineering and technology in order to achieve *"An economically strong and prosperous nation; acceptance at the highest scientific levels internationally; and a responsible, informed and responsive society."*

Whilst making a case for additional, strategic investment in science, engineering and technology, Batterham (2000) stresses the need for innovation in pursuing scientific advances and implementing them successfully. He argues that science, and the processes by which it is done, now requires active innovation in itself. It is no longer a question of whether change is needed – a continually search for improvement is essential for innovation. Maintaining the status quo is simply not an option (Batterham 2000). Research agencies, providers and funders must seek ways to continually improve in both scientific excellence and application.

Evaluation in light of the triple bottom line is a process for assessing, valuing and improving the investments in scientific research and development. It has been argued that evaluation that goes beyond the traditional economic impact assessment to contribute to institutional learning and performance can help agricultural research communities to respond to rapidly changing social, economic, political, technological and institutional conditions (Horton and Mackay 2003).

Accepting the premise that change and improvement is essential, this review explores the application and use of evaluation and the triple bottom line. It focuses on agricultural research, development and extension (RD&E), drawing from other fields where relevant. It does not investigate methods for evaluation but rather the methodologies that underpin a strategic, organisational approach to evaluation.

The Triple Bottom Line

Measuring global progress towards the goals of 'ecologically sustainable development' (World-Commission-on-Environment-and-Development 1987) was a key issue for the Rio Earth Summit. This summit, also known as 'Agenda 21', acknowledged that economic, environmental and social processes are interrelated and advocated indicators to measure these factors (UN 1992).

There is now an increasing trend and demand for organisations to demonstrate transparency and accountability beyond the domains for financial performance (CPA_Australia 2004). This sustainability reporting seeks to more closely align reporting, underlying management practices and measures of corporate performance with the ideals of sustainable development. Monitoring of the 'triple bottom line' (Elkington 1997) of economic, environmental and social performance has become embedded in arenas ranging from public policy to business performance. The triple bottom line (TBL) has been used as a process for integrating sustainability into the business environment, creating a new measure of corporate performance of balancing traditional economic goals with social and environmental issues (McDonough and Braungart 2002).

French President Jacques Chirac has led United Nations in discussions to expand to a "quadruple bottom line" for sustainable development that considers cultural diversity as the fourth pillar (UNEP 2002). This cultural diversity is regarded as a source of innovation, creativity and exchange.

Others have suggested that the fourth element of a quadruple bottom line is corporate governance. In his original discussion of the triple bottom line, Elkington (1997) viewed TBL implementation as the responsibility of the corporate board. He considered that TBL would lead to a better system of corporate governance that would lead to a chance for sustainable capitalism. Some organisations have taken this lead, building triple bottom line into their corporate governance structures and individual performance assessment (NovoNordisk 2002). However, investor concerns following major corporate collapses have triggered the formalisation of this by including 'corporate governance' as the fourth element of a "quadruple bottom line" (ANAO 2002; CPA_Australia 2004). This concept of governance adds accountability and transparency to the measures of outputs and outcomes. It considers whether all stakeholders are heard and their needs met in decision making.

Triple bottom line has now been widely adopted in Australia. Governments and many companies are required under legislation to report on some economic, environmental and social factors. TBL has been adopted more widely for voluntary reporting by organisations to demonstrate that they have sustainable growth – as both an internal management tool and an external reporting framework (Environment_Australia 2003). TBL frameworks can link with the mandatory environmental reporting requirements such as the National Pollutant Inventory (NPI) and State licensing regimes (Environment_Australia 2003). Major evaluations such as Australia State of the Environment Reports 1996 and 2001 (Australian_State_of_the_Environment_Committee 2001) and Country Matters: Social Atlas of Rural and Regional Australia (BRS 2001) provide benchmarks and identify achievements and challenges for the nation as a whole for environmental and social development.

Research investments are now also being assessed with regards to the triple bottom line. Seeking to demonstrate their true value, and with encouragement from the Commonwealth Department of Agriculture Fisheries and Forestry (DAFF), Australia's Rural Research and Development Corporations (RDC) are developing methodologies for reporting against triple bottom line outcomes in their annual reports to the Commonwealth (McGee, Charlton *et al.* 2002). Many are also realising the potential benefits that evaluation may give for investment planning, enhancing linkages across projects and focussing research on outcomes. The need for government agencies to be accountable in terms of the triple bottom lines was one of the triggers for the Victorian Department of Primary Industries (DPI) to build their evaluation capability (McDonald, Rogers *et al.* 2003).

Indicators of these TBL impacts are being developed by numerous groups including the CRDC (Roth 2003). Resource Condition Indicators are being developed to aid the measurement of performance of investments through national programs such as the National Action Plan and the Natural Heritage Trust (Monitoring-and-Evaluation-Working-Group 2004).

TBL RETURN ON R&D INVESTMENT

Whilst triple bottom line outcomes can be measured in a variety of ways, increasing pressure to demonstrate return on investment of public expenditure commonly leads to economic valuation processes being conducted. Various evaluation approaches, particularly Benefit-Cost Analyses focus primarily on measuring the accrual of economic benefits from project or program investments, in particular Benefit-Cost Analyses.

Identifying economic values for the social and environmental benefits has proven to be difficult as many of these outcomes expand over many years, flow on to other areas and frequently do not have clear market values (McGee, Charlton *et al.* 2002). As environmental and social impacts often involve benefits and costs that are not bought and sold in markets, market values will not identify people's preferences for these values (Dumsday and Sturgess 2003). Many of these social and environmental benefits can best be measured using qualitative techniques. However, accountability is increasingly requiring clear economic valuations of all returns.

This dilemma is clearly evident in an AFFA report (McGee, Charlton *et al.* 2002) that within a few paragraphs states: "Measuring social and environment projects for example is more difficult to quantify……lending themselves to more qualitative assessment" whilst in the previous paragraph stating that "While it is important to ensure that social and environmental benefits are captured, return on investment, measured through benefit-cost analysis (BCA) is important to demonstrate value to the stakeholders…..The ability to quantify outcomes is necessary for RDCs to measure the returns for funding invested."

Without effective non-market valuation, Benefit-Cost Analyses can bias towards investments that have a clear economic production focus with less value placed on public good. For example, an evaluation for the RIRDC of the Joint Venture Agroforestry Program considers benefits in sales of wood, projected sales of carbon credits, salinity, biodiversity protection and aesthetic values. The study found high economic benefits from the commercial returns of this program but little for the environmental and social benefits (Bauer, Kirchner *et al.* 2003).

Tools for non-market valuation

Tools and approaches are currently being developed to help place economic value on these non-market features. These tools include 'revealed preference' and 'stated preference' techniques (Dumsday and Sturgess 2003).

Revealed preference techniques rely on observations of people's actions in buying goods and services that are in some way related to the non-market impact. These include a travel-cost method (the relationship between people's preferences for recreational experiences such as natural areas or sites, and their willingness to pay the costs of travelling to the site) and 'hedonic pricing' (estimation of the price of a marketed good as related to its determining factors, including non-market elements).

Stated preference techniques use surveys to ask people questions that indicate the strength of their preferences. The most widely used stated preference technique is 'contingent valuation method' which uses a structured questionnaire to ask respondents their willingness to pay. 'Choice modelling', a further development of this technique asks respondents to select between alternative future scenarios that are described either by non-market attributes or by characteristics and a monetary cost.

Other techniques include 'benefit transfer' (using values that have been estimated in a previous study); a cost saving or relocation method; replacement cost; interpretation of similar decisions; preventative expenditure; and threshold analysis. These techniques can be used for valuation of environmental benefits and some social benefits can be addressed under environmental categories. Methods for monetary valuation of other social impacts are not yet well developed (Dumsday and Sturgess 2003).

Non-economic valuation of TBL

In light of the complexity involved in defining environmental and social values in monetary terms, alternate approaches, both qualitative and quantitative are also used to measure impact against the three bottom lines. Considering that evaluation is conducted primarily to aid decision making (Owen and Lambert 1998), factors other than economic return on investment that influence decisions are worth exploring.

For example, social statistics, published regularly by the Australian Bureau of Statistics since 1978 (www.abs.gov.au), provide a basis for comprehensive and balanced judgements about the conditions in a particular society (Armstrong and Francis 2003). These social indicators, such as crime statistics or employment rates, can also be highly influential in judgements, decisions, political direction and investments, particularly in the public arena. An overview of the numerous projects in Australian and elsewhere that investigate social and community indicators has been compiled (Salvaris 2000).

Numerous resources and guidelines provide advice for reporting on non-financial performance, including:

- Environment Australia's *A Framework for Public Environmental Reporting, an Australian Approach*-a step-by-step guide to reporting specifically tailored to meet the needs of Australian organisations. www.ea.gov.au/industry/finance/per.
- The Group of 100's Sustainability: *A Guide to Triple Bottom Line Reporting* - aimed at providing senior executives with a high level understanding of TBL reporting. www.group100.com.au.
- A statement from Commonwealth Government, business and community representatives on the value of Corporate Social Responsibility and TBL practices for Australia. This statement is an initiative of the Prime Minister's Community Business Partnership. www.partnership.zip.gov.au.
- *Guide to Social Indicators and Methodologies,* prepared by the Commonwealth Department of Family and Community Services www.facs.gov.au.
- *A Guide to Economic Indicators and Methodologies,* prepared by the Commonwealth Department of Family and Community Services www.facs.gov.au.
- Triple Bottom Line Reporting in Australia - A Guide to Reporting Against Environmental Indicators. Environment Australia. June 2003 www.deh.gov.au/industry/finance

DOES TBL WORK?

It has been argued that while TBL reporting has its value, there are fundamental problems with this approach to evaluation with regards to issues of social and ecological sustainability (Baker 2004).

McDonough and Braungart (2002) caution care that a business strategy focused solely on the bottom line, even if that is the triple bottom line, may be at risk of missing opportunities to pursue innovation and create value in business design. In a business context, there is no evidence to suggest that "socially conscious" companies are more profitable (Hayward 2002). Whilst several large businesses that have pursued social and environmental concerns have performed well, others have been significant failures and Hayward found no link between these goals.

A global survey of Chief Executive Officers found that whilst nearly 70 per cent of CEOs say that corporate social responsibility is 'vital' to profitability, they also consider that it has less to do with earnings and more to do with reputation across a broad array of stakeholders (ANAO 2002). One would expect that some market segments would be more responsive to demonstrated environmental and social outcomes than others. Some companies specifically target these markets.

Evaluation

"....evaluation is not a fault-finding exercise but rather a strategic management tool that allows management and programme staff not only to review progress made but also to critically reflect on the constraints and challenges of delivering a quality ... programme".

Dr. Klaus Topfer, UNEP Executive Director (UNEP 1999)

Providing a process for the measurement of progress towards the triple bottom line is one of the applications for evaluation. More broadly, evaluation is the systematic collection of data to aid in decision-making (UNEP 1999). By providing evidence and empirically based knowledge, evaluation has the capacity to enhance decision-making, to involve others along with senior managers in important decisions, and to enhance the effectiveness of organisations (Owen and Lambert 1998). It is used, amongst other things, to inform political decision-making and debate, to inform investment decisions and to improve programs.

Evaluation is an important tool in determining whether a program has achieved its objectives and whether the techniques used were the most effective and suitable. It is also a tool for learning and organisational development (PlanningNSW 2003). By identifying and communicating the impact of projects, evaluation can help to identify and overcome blockages between research and adoption (McDonald, Rogers *et al.* 2003). This approach was applied in Victoria where four reviews of the Agriculture Division of the Department of Primary Industries identified that the potential impact of the strong science capability of the division (described by one reviewer as Victoria's 'best kept secret') was not being fully realised (DNR&E 2001).

Evaluation is in no way restricted to agricultural research and development programs. Health and welfare departments, education institutions, defence forces, planning and international development agencies are just a few of the diversity of sectors that undertake evaluation of their programs and investments. Many of Australia's CRCs, spanning a wide diversity of disciplines, are actively seeking frameworks and approaches for evaluation of their impact, and in particular for valuation of public good outcomes (Rob Fearon, *Pers. Comm.* 2004). The overarching principles and approaches to evaluation that have been used in these other contexts can be adapted to research and development programs.

A review for RIRDC collated and assessed frameworks and approaches for evaluation in agricultural extension in Australia (Dart, Petheram *et al.* 1998). In the few years since then, much has progressed in evaluation in agricultural research and extension (Dart, *Pers. Comm.* 2003). Rural RDCs are looking for evaluation not only of extension programs but of whole research and extension programs.

PHILOSOPHIES AND PURPOSE FOR EVALUATION

The underlying philosophies about and purposes for evaluation vary between different groups. They will influence the way by which evaluation is approached, conducted, valued, trusted and used. One of the areas where philosophies commonly differ is the degree to which the approach is participatory or external. For example, it has been argued that outcome-oriented evaluations are well suited to 'transfer of technology' programs but are not appropriate for participatory extension approaches (Murray 2000) and presumably the same argument holds for participatory research programs.

Parallel to this are differing schools of thought on the purpose of evaluation: on the one extreme, a focus on assessing the economic impacts of research and on the other an approach of providing the research community with the information to improve their work (Horton and Mackay 2003). In practice, these could be considered a continuum and the literature shows that many approaches are a mixture of these purposes. It is worth being aware of which position along this continuum the various stakeholders come from.

There are multiple purposes for evaluation, which influence the manner in which it is approached. Some thoughts on the purpose for evaluation have been:

- Influential evaluations can change attitudes or behaviours, persuade others, justify policy and public expenditures, empower change agents, place an item on the public agenda and have numerous other effects (Mark and Henry 2004).
- "The ultimate purpose of evaluation is social betterment" (Henry 2000).
- "Accountability to industry and government stakeholders is an important feature of the Research and Development Corporation model" (McGee, Charlton *et al.* 2002).
- In research and development, evaluation can provide valuable feedback for on-going improvement of programs and activities, assist in targeting resources for greatest impact, be used to tailor programs to client needs and provide accountability to funding agencies and clients (UNEP 1999).
- The basic principle underlying all evaluation techniques is estimation of the costs and benefits of the activity in a way that facilitates comparison of results. (AACM-International 1997)
- The W.K. Kellogg Foundation outlines a strong philosophy of evaluation for "good thinking". Evaluation is embedded in projects from the outset and used, among other things, to improve processes, target to stakeholders needs and enhance learning within projects. (Kellogg-Foundation and Sanders 1998).
- Evaluation is to help projects become even better than they planned to be.... First and foremost, evaluation should support the project.... W. K. Kellogg Foundation, Evaluation Approach, 1997
- Understanding the outcomes of past projects is an essential element in planning research programs and in selecting projects for funding (Gordon and Davis 1999).

These views illustrate that evaluation is primarily driven by the perspectives of:

- Accountability – to demonstrate return on investment (which may be a legislative requirement);
- Identifying strategic direction and investment; and
- Learning and continuous improvement where project management and project evaluation are intertwined.

Table 1 expands on these major purposes for evaluation in agricultural research and development. Well designed evaluation programs can contribute to these multiple goals, thus providing greater value from the evaluation investment and the R&D investment overall.

Table 1: Purposes for evaluation of research and development (Templeton 2004).

Reason	Purpose
Accountability	To describe to stakeholders the research-induced impacts, evidence that outcomes have been achieved, estimated return on investment. Impact assessments are used to underpin private and/or public support for continued investment. Often at completion of a project.
Management decision-making tool	To better inform manager on complementarities and trade-offs between activities or projects within the investment portfolio. Aids to target limited R&D funds to projects with higher expected returns. Opportunity cost of investing in projects with lower returns (eg where economic return is not the main priority). Formative at project design and selection stage.
Increasing Awareness	To increase project team's awareness of the broader implications (or lack of) of their RD&E. Can be at various stages of project lifecycle. To help improve efficiency of current and future research.
Identifying research-to-impact pathway	To identify factors that influence uptake of research results. Focus on the users of the research outputs rather than the outputs themselves. - *Ex ante* to plan pathways for research to lead to impact. - M&E throughout project implementation gives timely information on whether research goals are being met and are readily available and weak links in the research-impact pathway. - *Ex-post* impact assessment identifies delivered outcomes and economic returns.

Another purpose, is to identify and communicate the impact of a project (McDonald, Rogers *et al.* 2003). Whilst this may fit within the 'Accountability' bracket, promoting positive results from research in the development new technologies or approaches can also be a valuable tool in the extension process. Knowledge about projects impacts can also be used for political agendas. For example, the United States Department of Agriculture (USDA) require all researchers to provide a one line "impact statement" they conveys what this project will mean to the broader community, environment or economy (Ellsworth 2002, *pers. com.*) Whilst many researchers are unhappy about this need to reduce their work into a single statement, these impacts statements are stored in a database that can be searched as needed for members of congress.

The use and purpose of evaluation will also vary throughout the stages of a project. *Ex-ante* evaluation is conducted prior to or at the start of a project to aid in project planning; *ex-post* evaluation follows the completion of a project to review outcomes, impacts and the process; and monitoring and evaluation (M&E) occurs throughout a project lifecycle. Five forms of evaluation have been described in a conceptual framework (Owen 1993) that considers the purpose of evaluation at different stages of a project:

Form 1 - Evaluation for impact assessment;

Form 2 - Evaluation for program management;

Form 3 - Process evaluation;

Form 4 - Evaluation for design clarification; and

Form 5 - Evaluation for program development.

These forms link the type of evaluation with the information that is needed for planning different stages of research activity (Figure 3). Other models present similar concepts in various forms (Horton 1998).

Rural research and development corporations have tended to focus their evaluation strategies on monitoring and evaluation of projects (Form 2). This is done largely through annual or biannual reporting requirements.

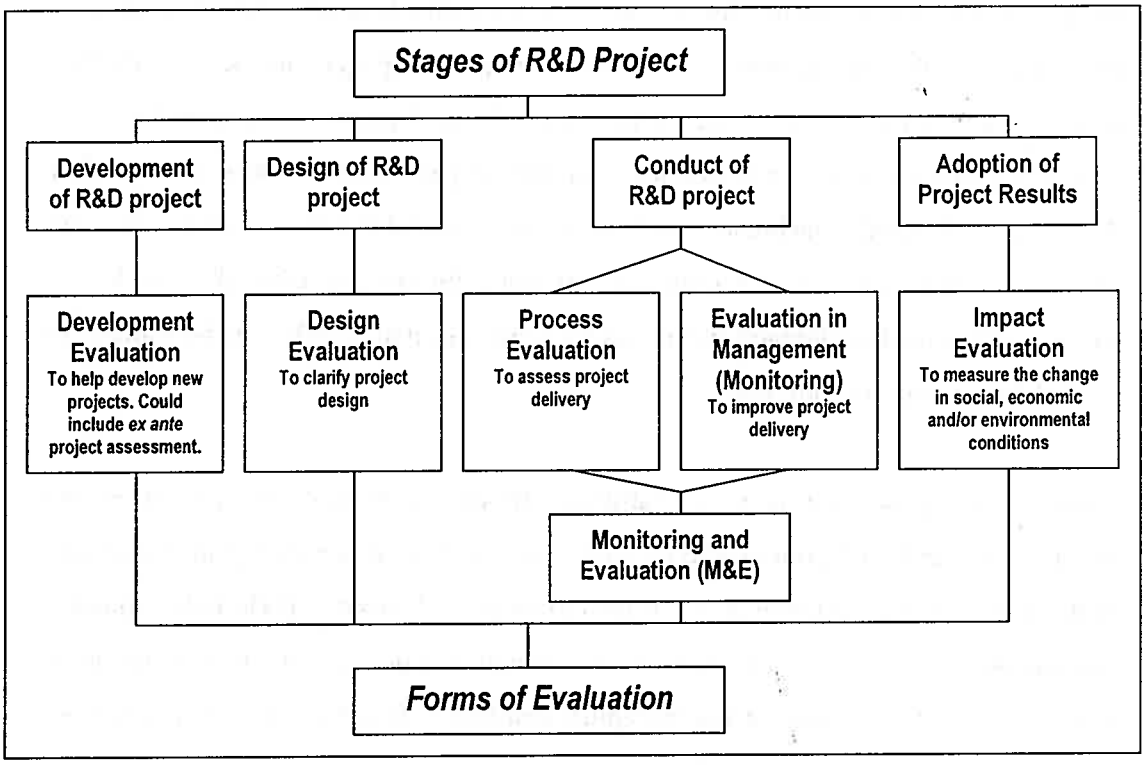


Figure 3 **Owen's Five forms of evaluation provides a conceptual framework based on stage of research activity.**
Source: (Templeton 2004) adapted from (Owen and Rogers 1999).

Perhaps aware of the differing philosophies on evaluation, many groups clearly define and communicate their evaluation principles. For example, the principles for the evaluation processes for Landcare programs are have been detailed in the National Natural Resource Management Monitoring and Evaluation Framework (NRM-Ministerial-Council 2003) as:

- is **useful for all partners** in natural resource management – Commonwealth, States and Territories, and regions, communities and industries.
- is **simple, cost-effective, affordable and practical**. To suit these ends the data infrastructure required to support the framework:
    - *avoids duplication of effort;*
    - *uses data for multiple purposes;*
    - *ensures that users can obtain the data*; and
    - *ensures that users can easily find out whether suitable data already exist.*
- recognises that **NRM interventions encompass a range of time-scales.**

- **supports meaningful interpretation of data over time** by establishing standard national indicators, protocols for their sampling, measurement and interpretation , and data quality and management requirements.
- **specifies the assumptions** on which monitoring and evaluation activities are undertaken in a consistent manner which is open to all stakeholders.

It is interesting to compare this with the guiding principles that were used by the project team for evaluation of the northern Farming Systems projects (Lawrence 2002):

- to be a natural part of each Farming Systems project;
- cover the life of the projects with regular feedback;
- be truly participatory and involve all stakeholders;
- develop a learning/reflective culture for improvement;
- consider the projects' processes;
- be simple; and
- maintain a balanced perspective that respects others' views.

Clearly the principles and philosophies vary with the context and people involved and so it is valuable to clearly identify and communicate these to team members and other stakeholders.

Terms of engagement

Evaluation requires some level of participation by or consultation with the key stakeholders. For participatory evaluation, stakeholder engagement is a key part of the evaluation process. Dick (1993) suggests that in the preliminary stages of an evaluation effort be invested in: building an effective and flexible working relationship between the evaluator and key stakeholders; identifying and involving other stakeholders in the process; and negotiating the role of all participants. If working with a representative group of stakeholders, he suggests setting up two-way communication means between them and the stakeholders they represent.

PLANNING EVALUATION

Evaluation strategies are most effective if they are tailored to the process or program in question (PlanningNSW 2003). This need not mean that it is essential to start from scratch in each case. This section describes the planning steps and frameworks that can be used in developing the evaluation strategy. Identifying guiding principles for evaluation, as described above, can aid to set a plan in a direction that engages all key stakeholders.

Developing an evaluation plan helps to clarify the evaluation not only for the evaluator but also for anyone else involved – team members, peers, funders and the group being evaluated. In planning the evaluation with stakeholders, it is useful to consider these steps (Wissemann 1992):

1. Clarify the Evaluation Request
2. Focus the Evaluation
3. Identify Information Needs
4. Identify Information Sources and data collection methods
5. Decide on the methods of data analysis and reporting procedures
6. Sequence evaluation activities and produce a written evaluation plan, including timeliness and assigned responsibilities.
7. Ensure stakeholders are satisfied with and support the evaluation.

Table 2 describes the key elements for a good framework or strategy for evaluation of community engagement initiatives (Qld-Government in press). These principles could be used as a plan and checklist in planning any evaluation strategy, particularly where the aim is to use the evaluation to influence change in practice or ownership of a program.

Evaluation methods can be used as a part of a learning or change management process. For example, empowerment evaluation (Fetterman 1996) involves participants self-evaluating. This can be an effective learning process to build into educational programs.

Table 2: Key elements of a good evaluation framework (Qld-Government in press)

- It has been developed through a participatory process
- It describes in detail the purpose of the evaluation and its intended use, including:
  - Who the audience is;
  - What they want to know;
  - When they want the information;
  - In what form they want it; and
  - How they will use it.
- It describes in detail the methodological approach of the evaluation, including:
  - The evaluation questions;
  - The performance criteria and indicators;
  - The type of data collected;
  - How data will be collected, analysed and interpreted; and
  - Who will be involved in the evaluation process.
- It has a clear timetable that is achievable and allows enough time for responses.
- The performance criteria selected reflect the agreed and documented:
  - Best practice principles or benchmarks; and
  - Intended outcomes of the activity.
- It includes an explanation of how the intended outcomes relate to the activity through a program logic model.
- It uses data collection methods that are appropriate to the research objectives and the research participants.
- It makes an effort to identify and understand unexpected impacts and outcomes of the activity.
- It analyses the context of the activity (including the political, social, economic and cultural contexts) in order to understand how this has affected the processes and outcomes of the activity.
- It includes a plan for sharing and using the learnings from the evaluation, including identifying what products/ forums (for example, reports, presentations, workshops or training) will result from the evaluation, and where relevant, how the evaluation will feed into a larger scale or meta-evaluation processes.

To evaluate or not?

> *"In this era of budget limitations and accountability, evaluation…has become almost as certain as death and taxes." (Brown 1979)*

In whatever manner it is done, undertaking evaluation requires some investment of time and resources. However, not every program, project or stage of a project needs to be evaluated. Careful thought is needed to determine the allocation of and best use of evaluation resources and the first step in evaluation planning is to determine whether there is benefit in evaluating that program or project.

In choosing whether or not to evaluate a program, (Boleman and Cummings 2002) suggest consideration of:

1) Will the gained information be useful?
2) Will evaluations be feasible?
3) Are the evaluations proper?
4) Are evaluations accurate?
5) Is the purpose/need clear?
6) Will the data be used?
7) Some one-time only programs may not need to be evaluated.

It will also be necessary to determine the boundaries of the particular program or project being evaluated, the context within which it operated and other projects or programs which may have influenced or contributed to the outcomes.

EVALUATION FRAMEWORKS

Frameworks for evaluation are useful tools for planning and reporting on both the evaluation and the projects themselves. The use of a uniform framework across an organisation or funding body provides a mechanism to compare and collate information about projects and programs in a consistent manner and for communication internally. Frameworks can help to provide some clarity about the many levels at which research and extension programs are targeted and thus can be evaluated. An existing evaluation framework may be adopted provided it is tailored for use in each situation (PlanningNSW 2003).

There are numerous frameworks that are used for evaluation of programs and projects. Dart and her colleagues (1998) identified many of these that have been used in agricultural extension. Here I will not attempt to review all frameworks but rather to present a few that may be useful in the context of cotton research programs.

Program Logic

Program logic or program theory is a description of the events of a program that lead to the ultimate outcome (Dart, Petheram *et al.* 1998; Scriven 1991) through a logical chain of relationships from inputs – outputs – outcomes – impacts. This logic may be drawn as a visual diagram. It is well suited for use in a team environment to question the assumptions between the various steps and the realistic outcomes and process for a project. Program logic models are ideally used from the early stages of project planning right through the life of the project. The most widely used and accepted frameworks in agricultural RD&E are mostly variations of a logical framework.

Logic Models

The simplest and clearest form of these is the logic model. Logic models provide a simple framework that can be applied to a small program, a process or a large organisation to represent the program's "theory of action" (Taylor-Powell 2001). Quite important to the value of the logic model is the definition of the assumptions (beliefs about the program and how it is thought to work; guiding principles) and the context (external conditions in which it exists and which influence the program success). This logic model has been further developed and adopted by the University of Wisconsin Cooperative Extension Service (Figure 4).

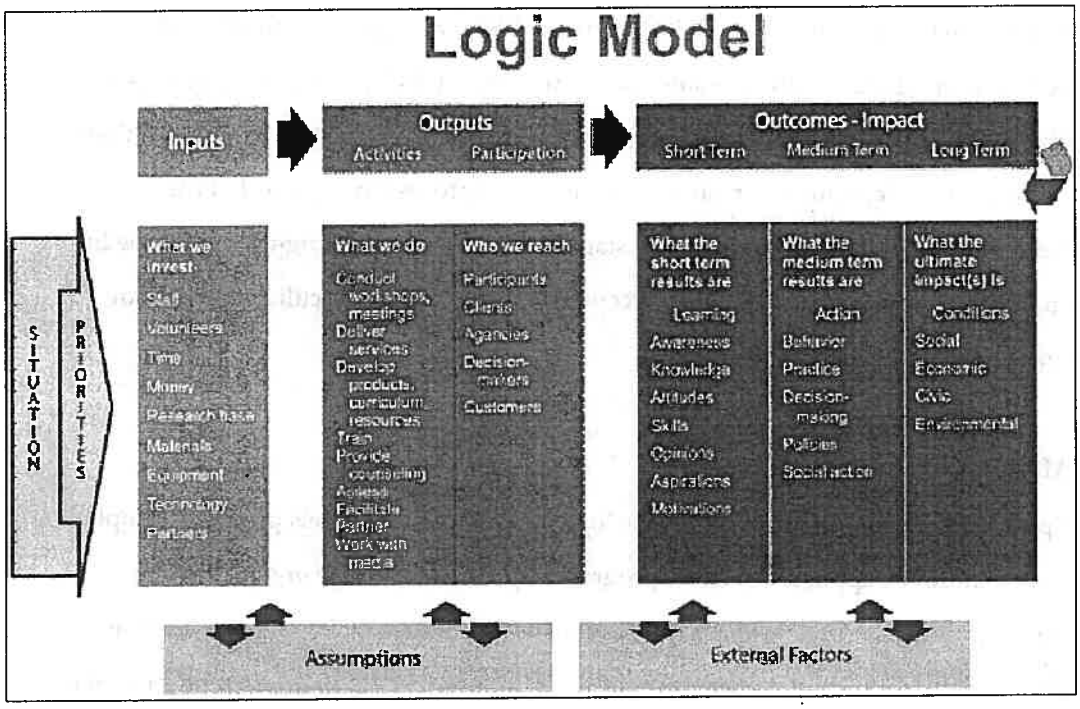


Figure 4 **Logic Model for evaluation as used by the University of Wisconsin Cooperative Extension (Taylor-Powell 2001).**

Bennett's Hierarchy

This hierarchy was developed almost 30 years ago for evaluation of extension programs in the USDA (Bennett 1977). The framework, derived from a logic model, is now widely used by many extension practitioners and agencies, including the Victorian Department of Primary Industries. This framework is specifically developed for extension programs. In consultation with Claude Bennett, DPI Victoria have modified Bennett's Hierarchy for use with research programs. The model can be used both in planning and in evaluating extension programs (Rockwell and Bennett 1995) as depicted in Figure 5.

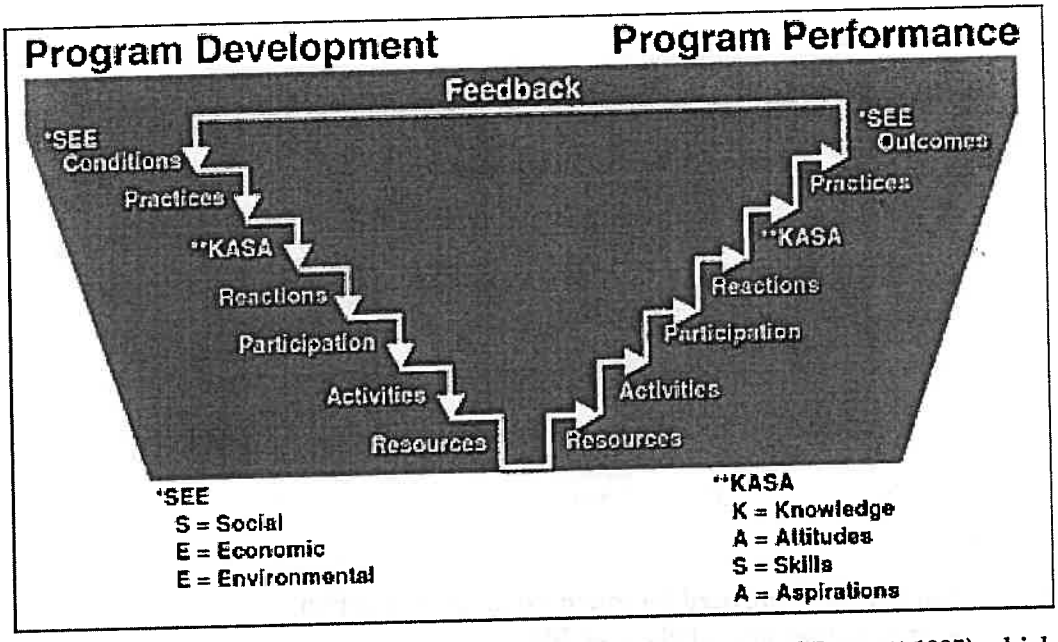


Figure 5 **Targeting Outcomes of Project model (Rockwell and Bennett 1995) which is an adaptation of Bennett's original hierarchy (Bennett 1977).**

Cotton Extension

A slightly modified Logic Model that separates clearly Outcomes from Impacts, and uses Bennett's Hierarchy (Bennett 1977) to help define the logic, has been introduced to the cotton extension program (Figure 6). The simplicity of the logic model makes it readily applicable across all programs and allows great flexibility in choosing evaluation methods to suit the issues and the evaluator (Christiansen, Pyke *et al.* 2003). The detail of Bennett's Hierarchy can then be used for greater description and planning where appropriate. The logic model matches closely with the manner in which CRDC, the primary funder of cotton extension, reports to the Commonwealth.

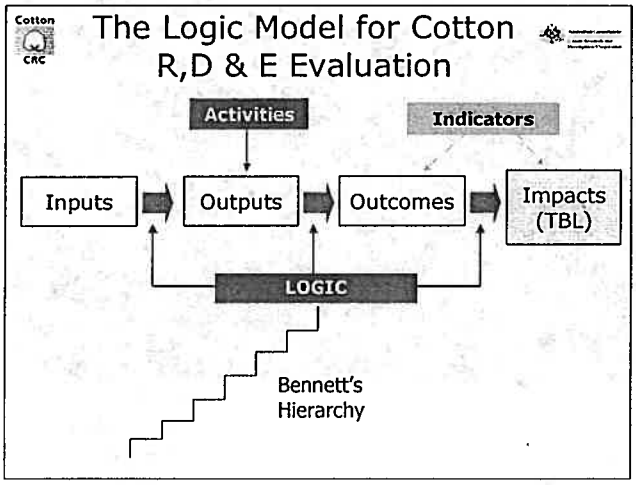


Figure 6 **Framework being tried for cotton extension evaluation**
Source: Christiansen, Pyke et al. 2003.

Logical Framework (Logframes)

The logical framework (matrix) helps to identify the essential elements of a project and the linkages between these in a logical, concise and objective manner, placing the project in the larger context (Dart, Petheram *et al.* 1998). In a more structured manner than other logical models, the logframe helps users to identify the inputs, assumptions and indicators for M&E. As with other logic models it can be used both in planning and to review and report on progress. The logframe works on the principle that each level of the hierarchy is needed for the next level to be achieved.

Logframes are used extensively in overseas agricultural projects (Dart, Petheram *et al.* 1998), including those funded by AusAid and locally by Land and Water Australia and the National Program for Sustainable Irrigation. There is a risk that researchers and other practitioners, particularly those new to evaluation, can find logframes to be overly prescriptive and onerous. Table 3 details the components of the logframe.

Table 3 Questions to be answered in using the logical framework.

Source: Dart, Petherham, et al 1998; modified from Farrington and Nelson (1997)

	Narrative summary	Measurable Indicators	Means of verification	Assumptions
Broader Goal	What wider problems will the project help to resolve?	What are the measures to tell whether the problems in cell A2 have been resolved?	What are the sources of information? What methods should be used for obtaining it?	(Goal to supergoal) What external factors are needed to sustain objectives in the long run?
Purpose	What immediate effects are intended for the project area or target group? What benefits are expected? What improvements or changes will the project bring about?	What are the measures to judge the project's immediate effects, benefits and losses?	What are the sources of information? What methods should be used for obtaining/ accessing it?	(Purpose to Goal) If the project achieves its Purpose, what external factors must be true if it is to help reach the goal?
Outputs	What outputs will the project produce so it can achieve its purpose?	What kind and quantity of outputs and by when will they be produced?	What are the sources of information? What methods should be used for obtaining/accessing it?	(Output to Purpose) If the project produces the outputs, what factors outside its control might hamper the achievement of the purpose?
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? What methods should be used for obtaining/accessing it?	(Activity to Output) What external factors must be realised to produce the planned outputs in time? What actions outside the control of the donor are necessary to begin the project?

Definitions used in this framework are (Dart, Petheram *et al.* 1998):

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.

Outcomes Hierarchy

An outcomes hierarchy aims to make explicit how a program is expected to work. By doing this, it can guide the planning and evaluation of activities. Cause and effect chains are identified that lead from the immediate intended project outcomes to the desired broader outcomes. An example of an outcomes hierarchy used in a Queensland Government engagement activity is provided in Figure 7.

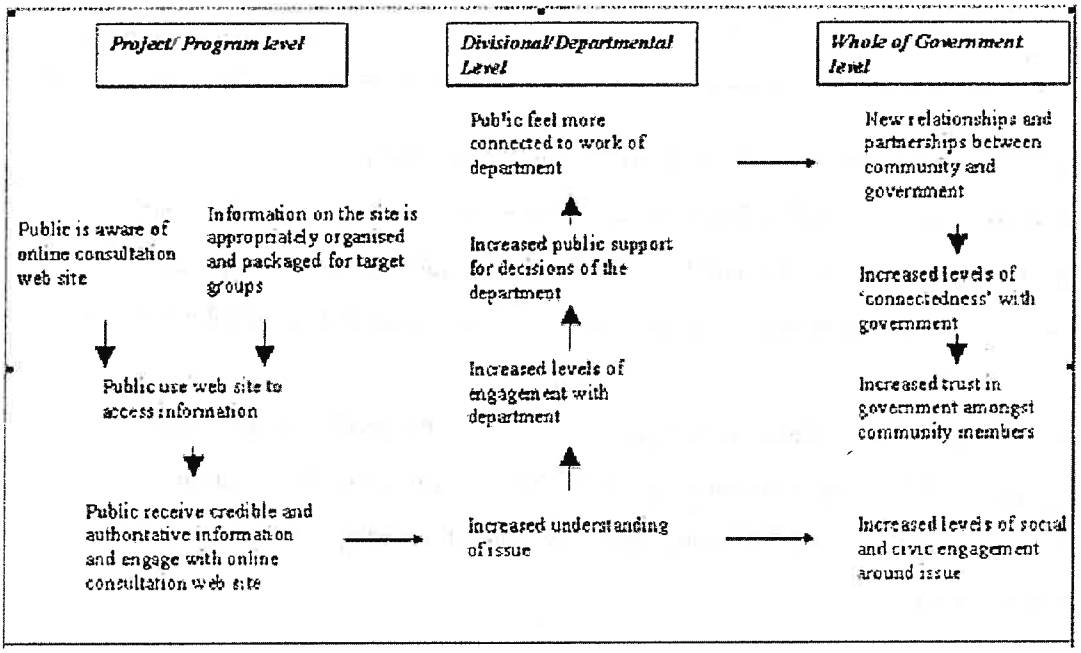


Figure 7 **Example of an outcomes hierarchy for a community engagement online consultation website project.**
Source: Community Engagement Evaluation Strategy (GetInvolved 2004)

Benefit Cost Analysis

The benefit cost analysis (BCA) techniques were developed for use by public agencies in evaluating the welfare or net social benefits accruing from large-scale public investment programs. The use of the BCA in rural systems has focussed on economic assessment of impacts. The basic BCA framework can also be used for non-economic measures as described in Table 4. AACM-International (1997) go so far as to argue that all evaluation techniques should be a variation of the benefit-cost analysis framework.

BCA can be applied *ex ante* or *ex post*. Sloane (1999), suggests that BCA of technology transfer and the adoption process in Land and Water Australia is best applied *ex post*. The challenges associated with utilising BCA for economic evaluation of triple bottom line outcomes are described in an earlier section.

Table 4 Types of Benefit Cost Analysis (AACM-International 1997)

Descriptive BCA: Social surveys, qualitative measurement of beneficiary perception and outcomes;

BCA: Included quantified benefits;

BCA (Multi-criteria analysis): explicitly includes other criteria such as non-market benefits.

BCA (Investment Decision Analysis): explicitly includes risk

Snyder Evaluation Process

The logical and BCA frameworks are essentially linear models based on the logic that, if programs are effective, inputs flow through to impacts. There is growing belief, well documented in the extension and international development literature, that the linear, technology transfer view is not as effective as participatory change models. The Snyder process (Dick 1993) meets a compromise by a participative, action research approach to the logic of resources ⇒ activities ⇒ effects ⇒ objectives ⇒ vision.

It aims to seek the involvement of all stakeholders or their representatives, encouraging their critical reflection within a flexible, systematic process. The process combines three main components, each of which builds on the previous by using a different form of evaluation: 1) Process evaluation to develop a better understanding of the functioning of the unit, including linkages, goals and contribution towards a vision; 2) Outcome evaluation – participants assess which of the goals have been achieved and how well this is done. Performance indicators can be set-up for on-going feedback and monitoring; and 3) Short cycle evaluation applies phases 1) and 2) to give regular feedback about performance which can be used to continually improve.

This was the overall approach that was used for building evaluation capability in the northern Farming Systems projects (Lawrence, Carberry *et al.* 2001).

Action Learning Cycle

Kolb's (1984) action learning cycle of Plan-Act-Review-Reflect can be used as the basis for an evaluation process. Among other things, this cycle is used as the basis of Environmental Management Systems including the Cotton Industries Best Management Practices Program (Williams and Williams 2000).

The model can also be used as an evaluation frame to identify the impact of a program at each stage. Figure 8 presents an evaluation process that can be linked to the action learning cycle in this way. The reflective process can also contribute to learning by project team members (Lawrence *pers. comm.* 2004).

The evaluation framework for the northern Farming Systems Projects was based on a combination of the action learning cycle and the traditional scientific approach (Lawrence, Christodoulou *et al.* 2003) that was developed from the Snyder approach (Dick 1997). In this project, evaluation plan and evaluation workbook were developed and used by many members of the project teams. Particularly valuable to the process was the clear definition of key questions for the evaluation that were developed by project teams members early in the life of the project.

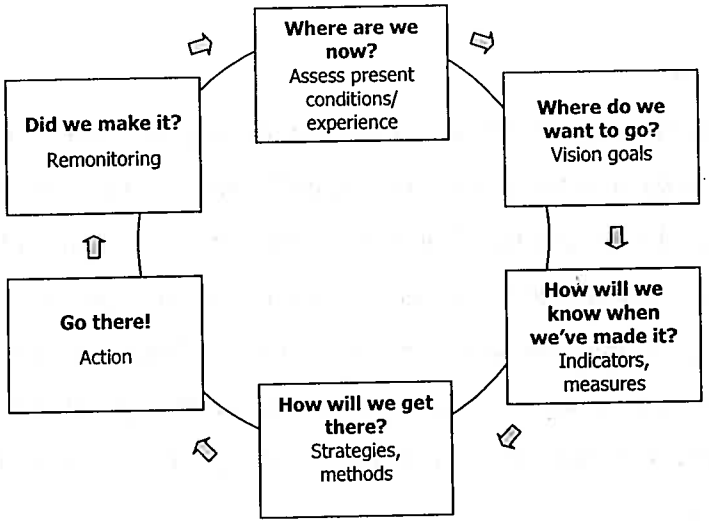


Figure 8 **The consumer participation planning model has similarity to Kolb's action learning cycle.**
Source: (National-Resource-Centre-for-Consumer-Participation-in-Health)

Sustainable Livelihoods

Sustainable livelihoods approaches pose an alternative model for approaching program planning and evaluation. Sustainable livelihoods is a people centred approach, developed as a process for use in development programs aimed towards reducing poverty (Carney 2002). The guiding principles of the Sustainable Livelihoods approach are: People Centred; Empowering; Responsive and participatory; and Sustainable. The principles of applying the sustainable livelihoods approach are: Multi-level and holistic; Conducted in partnership; Disaggregated and Long-term and flexible. Sustainable Livelihoods thinking can be used in designing an evaluation process, and as a checklist for monitoring.

The Sustainable Livelihoods framework has been used to assess the fit and the impact of agricultural research on poverty in developing countries (Adato and Meinzen-Dick 2002). It is a multi-disciplinary approach that uses both qualitative and quantitative methods.

Balanced Scorecard

The Balanced Scorecard was developed as a performance management tool in business (Kaplan and Norton 1992). This tool gives a fast, comprehensive overview of financial measures complemented with operational measures on customer satisfaction, internal business and innovation and learning that affect the current and future performance of the organisation (Figure 9). It aims to balance long and short term objectives using quantitative and qualitative performance measures that are viewed from different perspectives of performance. Kaplan and Norton (1996) stress that it is a template only and needs to be customised to each situation.

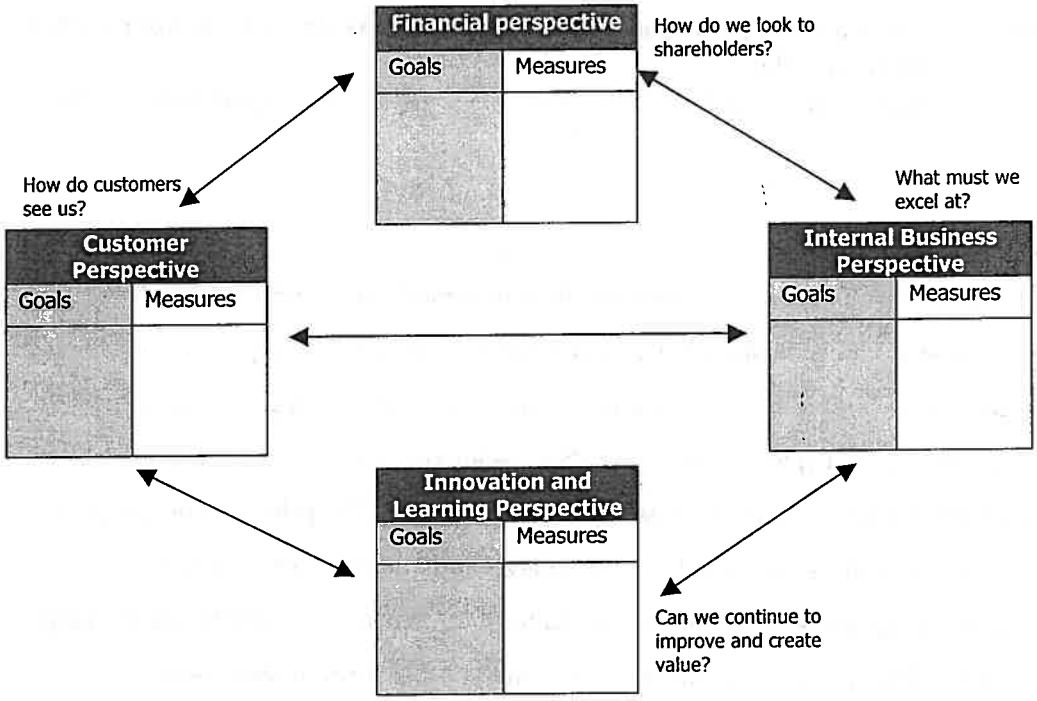


Figure 9 **The Balanced Scorecard model to link performance measures.**
Source: (Kaplan and Norton 1992)

The Balanced Scorecard (BSC) model has been used to evaluate Information Technology and Information System (IT/IS) investments. Stewart and Mohamed (2001) found that the original BSC was too narrow a way to look at the evaluation of IT implementation in construction. They proposed a further development of a tiered BSC for this purpose.

There may be potential to adapt the balanced scorecard model for use in evaluation of agricultural research investments. Tailored to suit a specific R&D organisation or program it may provide an alternative manner for reporting triple bottom line outcomes that doesn't rely on economic valuation of all features.

PROGRAMS OR PROJECTS?

As research *programs* are comprised of *projects*, it will be essential for the evaluation approach to include both. The frameworks and planning tools described can be applied to either a project or a program basis. RDCs and other organisations have used different approaches to the project-program basis for evaluation (AACM-International 1997). International development organisations such as FAO that have traditionally focussed on monitoring and evaluation of projects are shifting towards more program based evaluation (Dixon, *pers. comm.* 2003).

Measures of actual impacts are needed to clearly demonstrate return on investment but they can lose their value if too many assumptions are needed to attribute change to research and extension inputs. If program outcomes are determined by combining project evaluations, care needs to be taken to avoid double accounting (AACM-International 1997). Ekboir (2003) argues that, due to the high number of external influences, it is unrealistic to use impact analyses, or impact indicators, to evaluate the impact of research investments. The problem is further complicated by the timelag that often occurs before impacts appear. On this basis, it is recommended that research impacts be analysed in a more complex adaptive systems framework that considers external forces (eg markets, policy), the direct and indirect interactions amongst the players (researchers, input suppliers, farmers) and the nature and evolution of the technology (Ekboir 2003).

Lower level evaluations do not measure actual impacts but provide short learning loops that can give valuable feedback for enhancing project activities. They can also contribute to an understanding of how activities may have influenced impacts. In practice, evaluations at different levels are useful for different purposes (Christiansen, Pyke *et al.* 2003). Evaluation frameworks can be used to bring together evaluations of projects and programs.

The decision of how to balance the program-project focus of evaluation will vary depending on the organisational structures and funding arrangements. For example, the Victorian DPI have consolidated research from many (2500), small projects into fewer (60), big-budget projects. These 'projects' then form a logical basis for evaluation. Similarly, GRDC have moved towards funding large, multi-agency projects such as the Central Queensland, Eastern and Western Farming Systems projects which have also been evaluated (Lawrence, Christodoulou *et al.* 2003).

RDCs and CRCs will also need to be aware of the evaluation frameworks, if any, being used by their partner agencies or other funders. Different projects within a program are commonly funded from a few sources, each of which may have a different evaluation framework. For example, in the Cotton CRC two new environmental extension positions funded by the Qld Agriculture State Investment Plan (AgSIP) will join a team with two similar positions funded by the Department of Agriculture Fisheries and Forestry. The staff will be members of the Cotton CRC Extension Network (funded largely by CRDC) and the Environment Program of the Cotton CRC, joining staff funded from several sources and employed through several agencies. The AgSIP and DAFF funding bodies and the Cotton CRC (together with CRDC) all use different evaluation frameworks. At this stage the host agencies contracted to deliver projects do not use their own evaluation frameworks. The Queensland Department of Primary Industries and Fisheries (DPI&F) choose instead to report on the requirements of the funding bodies (Gus Hamilton, *pers comm.).* This may change soon with the new DPI&F structure articulating a need for evaluation at each level.

EVALUATION USE

Does the evaluation itself deliver a worthwhile impact? As discussed, there are a number of purposes for undertaking evaluation that clearly have potential to be beneficial to the research process and ultimately to social, economic and/or environmental betterment. In applying the evaluation principle to evaluation itself, does evaluation deliver a return on investment? Have the evaluations had an impact, either through stakeholder participation in the process or use of an evaluation report? With all the noblest goals, there is potential for the evaluation itself to be a wasted effort if it is not used and enacted. Use has long been a concern for evaluators (Grasso 2003; Patton 1978). Patton (1997) considers that 'doing evaluations that are useful and are actually used' remains the central challenge to the evaluation profession. These challenges relate both to the design of the evaluation studies and to the nature of decision-making, which research shows is not entirely rationally based (Mackay and Horton 2003).

Impact assessment studies have been suggested to fall along a continuum with academic research at one extreme and utilisation-focussed evaluation at the other (Mackay and Horton 2003). These authors go on to logically argue that impact assessment studies conducted in a manner towards the utilisation-focussed end of the continuum are likely to have greater use by decision makers.

Patton (1997) stresses the importance of the personal factor in evaluation – he recommends that the first step in a user-focussed evaluation is to identify all the stakeholders (potential beneficiaries) of the evaluation. Evaluators can improve the usefulness of their work by taking into account the multiple audiences they need to reach, the diverse needs of each of these audiences and the timeliness of the work (Grasso 2003). An example relating to the US Medicare system, illustrates that timeliness is so crucial in the usefulness of evaluation that a complete scientific approach may in fact delay the findings such as to miss the critical period for decision making (Grob 2003).

A quest for scientific "proof" will generally increase the costs of an evaluation but does not necessarily increase the usefulness for its purpose (Grob 2003). Grob (2003) suggests that quickly done, indicative evaluations that may be less scientific, less proven with tentative conclusions can be highly useful for decision making provided that the limitations are clear. He further encourages evaluators to engage users in the evaluation process.

The evaluation process itself can also deliver results. Evaluation was a core component of the Central Queensland Farming Systems Project and was helpful in developing project strategies and team-shared views of the project's direction and purpose. Early in the life of the project, team members collaboratively identified the key measures to evaluate the project. This process voiced differing perspectives of the goals and abilities of the project, which in turn helped the team to clarify these perspectives and develop team based strategies for this farming systems research and extension project (Lawrence *pers comm.* 2004).

BARRIERS TO EFFECTIVE EVALUATION

A number of potential barriers to evaluation will need to be managed for it to be a useful tool in research and development programs.

Lack of information

A lack of suitable information from projects and programs will increase the difficulty and complexity of the evaluation or make it less valuable. A cost-benefit analysis of 250 projects in RIRDCs emerging industries program was unable to indicate a value for 50% of the projects as 20% of the projects had insufficient information to make a judgement and a further 30% were inconclusive (Gordon 1998). For many other projects in that portfolio, outcomes were difficult to formally evaluate. Similarly, the USA's Office of Management and Budget found that half of the programs that they rated in preparing the USA 2004 budget were unable to demonstrate results (Kingsbury 2003).

Perceptions

Evaluation, as a form of social research is essentially a value-laden process (English and Kaleveld 2003). More technically focussed research scientists are often sceptical about such social research, particularly qualitative studies.

Evaluation may sometimes be regarded as a threatening process and has been seen to have the potential to undo or undermine an effective program (Brown 1979). For example, the beliefs about evaluation held by staff of the US Student Affairs Office (Patrick and Niles 1988) included:

- The evaluation process was 'value-loaded' and 'in competition' with other academic units for scarce resources;
- Program staff who have limited control over budgetary matters were concerned that negative evaluation findings could be used as justification for reduced funding;
- Programs may be developed haphazardly, with ill-defined objectives, making evaluation less likely to reflect the true nature of their effectiveness;
- Many of these professionals do not believe that their efforts are measurable, resulting in a reluctance to support an evaluation.

Politics

Evaluation occurs with an often highly politicised context. English and Kaleveld (2003) advise evaluators to acknowledge that evaluation will be highly politicised where stakeholders interests are threatened, to find a framework to understand the various political agendas, to use an agreed and explicit process and to be aware of their own interests and expectations.

Time and resources

As with any initiative, allocating sufficient time and resources to evaluation can be a challenge. This can be particularly problematic if it is expected that existing staff will take on the evaluation role without specific time allocation.

The hand that feeds...

Contradictions arise where an independent reviewer is funded to do an evaluation job by the particular agency in question (Scriven 1996), as commonly occurs. It is well recognised that the rewards for an evaluator in producing a favourable report often greatly outweigh those for an unfavourable report (English and Kaleveld 2003).

Current Status in Australian Agricultural RD&E

The main mechanism used by RDCs has been the annual and project reports that service monitoring and evaluation purposes (Templeton 2004). Most also have various types of reference groups comprised of key stakeholders who review research progress and provide advice on project investments.

In recent years, several RDCs and Cooperative Research Centres (CRC) have contracted external evaluators to undertake benefit-cost assessments. Several of these have been done through focussing on selected projects or program areas to calculate a value across the entire research portfolio. Rather than undertake studies of their entire portfolio, the Rural Industries Research and Development Corporation (RIRDC) has chosen to commission studies that focus on the value of specific clusters of projects, such as the Emerging New Industries (Gordon 1998), Prospective New Industries (CIE 2000) and the Joint Venture Agroforestry (Bauer, Kirchner *et al.* 2003) Programs.

In 2000-01, the Grains Research and Development Corporation commissioned CIE and Hassall & Associates to conduct an economic analysis of the entire GRDC investment portfolio (McGee, Charlton *et al.* 2002). Models developed in this study included Value Chain, Benefit Cost (Financial), Optimal Allocation and Systematic (Stochastic) Risk and Sensitivity Analysis. In the same year, GRDC also undertook *ex-ante* economic evaluation of it current (2000-01) and potential research portfolio (GRDC 2002), one of the few published *ex ante* evaluations of agricultural R&D at a strategic or portfolio level in Australia. The triple bottom line has been considered in some more recent assessments, such as the evaluation of GRDC's Farming Systems projects (Hassall&Associates 2004).

On the recommendation of a review by URS Australia, Land and Water Australia, use logframes for evaluation of their projects. Project leaders are required to submit a completed logframe and some projects in the portfolio are then selected for external evaluation, frequently by URS.

The Cooperative Research Centres program was established with the aim of strengthening collaborative research links between industry, research organisations, educational institutions and relevant government agencies in order to gain greater benefit from Australia's investment in R&D (www.crca.asn.au). Within this model there is scope to enhance the capability for evaluation that investigates both the value to industry and the public of the research investments and the value of the collaborative model. Several CRCs, including the Cotton CRC, have commissioned benefit-cost analyses of their entire investment portfolios.

CRDC regularly conduct external reviews of their programs. The review panels generally involve an international specialist in that research area together with a cotton grower who is involved with the Australian Cotton Growers Research Association (ACGRA). Recent examples include reviews of Insecticide Resistance research and management program and the Plant breeding and Biotechnology program. A review of the extension program is planned for 2005. These reviews are used primarily for planning the future direction and processes of the programs, and related investments (Owen's form 5).

Capability for Evaluation in RD&E

In reviewing the experiences of five, diverse USA public agencies, Kingsbury (2003) found the key elements of evaluation capacity were: an evaluation culture, data quality, analytic expertise and collaborative partnerships. For the purpose of this review, Kingsbury's 'analytic expertise' or skills component is regarded as *capacity*, which together with the other elements forms the evaluation *capability*.

From experiences in the Victorian DPI, McDonald, Rogers *et al.* (2003) offer seven key recommendations for building evaluation capability in public sector organisations:

- Start small and grow;
- Address both supply and demand;
- Work top-down and bottom-up;
- Use a theory of change behaviour;
- Develop a common evaluation framework, including a generic program theory;
- Build knowledge of what works within the agency's context; and
- Systematically and visibly evaluate each stage.

RD&E agencies are rich with skills in scientific process and adoption strategies. Building an evaluation capability can be aligned to draw on these skills and processes (McDonald, Rogers *et al.* 2003). The scientific research and evaluation capability building processes have in common: the processes of formally reviewing others work, trialling new techniques, inviting input from external reviewers, and visiting researchers at critical stages. Similar to extension process, building evaluation capability involves developing resource materials; developing ways to communicate information about available technologies; underpinning the message with the need for adaptation to issues and local conditions; and generating new knowledge.

The most comprehensive approach to building evaluation capability in an agricultural R&D organisation in Australia seems to be the case of the Victorian DPI. In that case, evaluation capability has been built gradually, with distinct, clear investment and purpose. Evaluation plans are now required with all new project applications. This has led to many research and extension staff thinking more critically about how projects achieve results; questioning the logic and rationale of their actions; improving what they do; and increased success in attracting external funds (McDonald, Rogers *et al.* 2003). Table 5 details the phases, inputs and strategies used in building evaluation capability in that agency.

Table 5 Phases, inputs and strategies used in building an evaluation capability in DPI Victoria (McDonald, Rogers *et al.* 2003).

	Phase 1: Addressing a specific need	Phase 2: Experimenting with volunteer projects	Phase 3: Mandatory evaluation for all projects	Phase 4: Expansion and consolidation
Number of RD&E projects involved	1	10	All new projects	All new projects
Time	18 months	18 months	2 years	Now and ongoing
Staff	1 person	2 people + consultants	2 people + consultants	4 permanent positions + 2 temporary
Role of the evaluation support team	Evaluator; researcher	Mentor; informal trainer; researcher	Mentor; formal trainer; researcher; facilitator; special events manager	As for phase 3 but with a stronger evaluation-research role
Evaluation of work on building evaluation capability	Independent blind review of reports	External evaluation using semi-structured interviews	Internal evaluation; mixed methods	Planned external review; continued internal evaluation
Research and development	Designed frameworks; experimented with realistic evaluation	Supported a PhD (Dart 2000)	Undertook research on: a framework for impact evaluation; social impact assessment; and values inquiry	Strategy for expanding evaluation research being developed

Some of the key elements of this program were (McDonald, Rogers *et al.* 2003):

- It was an action learning process of trial-reflect-grow;
- Researchers were originally left with open options to develop their own frameworks but this took too much time and became too complex so Bennett's Hierarchy was introduced.
- Short learning loops enabled evaluation to give quick feedback that was used to improve projects; it also gave project staff a first hand appreciation of the benefits of evaluation.
- External validation of the process and reflection by participants.
- Demonstrated value of the evaluation process was identified by and communicated to project staff.
- A member of the Target 10 project team that was the initial evaluation pilot study moved to a senior management position in the agency where he gave great support for building evaluation capability.

- An evaluation support unit that developed frameworks, delivered training to approximately 18% of staff of the DPI Agriculture Division and some from other divisions, provided support to project teams and did evaluation research.
- A mixture of internal evaluation with some external consultants.
- Building both supply and demand for evaluation capability.
- Trials, research and evaluation of the evaluation capability building process developed knowledge of what worked for that system, and helped to build trust and credibility with the research and extension staff.

DPI&F initiated an evaluation approach some years ago. Training courses in evaluation were conducted to build evaluation capacity of staff and an evaluation unit was established with specialised skills to deliver training and provide evaluation services (Wissemann 2004 *pers. comm.*). In this case the evaluation capability building strategy was less clearly planned and has not become strongly embedded in the organisation's practice.

In the case of GRDC's northern Farming Systems projects, GRDC funded DPI&F to deliver a complementary project with an extension evaluation specialist to support the evaluation of both the processes and impacts of the projects. This project set out to evaluate the impact of the farming systems projects on clients' learning, decision making and outcomes; develop an understanding of the participatory RD&E processes being used; and develop principles and procedures for planning and evaluating farming systems RD&E (Lawrence 2002). Acknowledging that many team members had limited experience with participatory RD&E process, evaluation was used as a process for reflection and learning within the project team. Over 50% of team members were involved.

Evaluation principles and procedures were developed in collaboration with members of the farming systems project team (Lawrence 2002), including:

- Guiding principles for evaluation of the northern Farming Systems projects;
- An impact matrix to satisfy the leaders of each project, the host agencies and the funders;
- A reporting framework;

- An evaluation package, including a workbook, with processes and tools to apply the above three issues and to integrate evaluation into planning and conduct of activities;
- A checklist for planning and managing farming systems projects (Lawrence, Carberry *et al.* 2001).

This approach has contributed towards team leaders developing both technical and process understanding from their RD&E activities and overall a more evaluative culture. Lawrence (2002) considers that this level of evaluation for learning and accountability would not have been possible without the evaluation project and reasons that ongoing evaluation support is required.

INVESTMENT

The level of investment to be allocated to evaluation will vary and there is no universal standard. AACM International (1997) suggest that allocating between 0.5 and 3% of funding towards evaluation can be justified economically, though usually no more than 1%. Some, slightly dated, examples of the proportion allocated towards evaluation include World Bank projects – 0.6-1%; ACIAR 0.7% (Lubulwa 1996).

McDonald (2003) and her co-authors suggest starting small to trial evaluation approaches and growing it across the organisation. Similarly, the CRDC have aimed to gradually grow evaluation skills and approaches in their RD&E programs (Pyke 2003 *pers. comm.*).

CULTURE

Building evaluation into the 'culture' of R&D organisations requires a long term view – at least 8 to 10 years has been suggested (McAllister 1998). The creation of an evaluation culture in an organisation is a step towards meaningful and useful evaluation. It has been argued to lead to positive change in the mix of 'working knowledge' used in problem solving and decision-making (Owen 2003).

An evaluation culture, which reviews achievements and processes to improve and target programs could be likened to a continuous improvement culture. A large body of separate literature expands on the concepts of continuous improvement and innovation.

Introducing mandatory evaluation alone is unlikely to create a culture supportive of evaluation. In fact, it may lead to token compliance, unreliable information and a risk of making the program even less relevant (Davies 1999). However, voluntary adoption alone is much slower and a mixture of approaches that builds towards mandatory evaluation has proven successful (McDonald, Rogers *et al.* 2003). Our experience in the cotton extension program is that a scheduled workshop session for extension staff to present evaluation results to their peers or a specific evaluation data gathering workshop has resulted in a far greater number and distribution of evaluation studies being conducted and reported (Christiansen 2003).

Some RDCs are attempting to influence research culture through their project application and reporting processes. For example, some require outcomes based project applications or include an evaluation frame in the application form. CRDC now includes questions about impact in their project report format. CRDC's extension reporting format was modified in collaboration with extension staff and staff training to base it on the logic model. Research in the sugar industry identified a need to build capability for evaluation and for a cultural change within the research community (Henderson 2003). The Sugar Research and Development Corporation (SRDC) undertook a few key activities to support their new evaluation approach (Henderson 2003 *pers comm.*), including:

1. Employed a research manager with a PhD and prior experience in cost-benefit analysis and other evaluation methods;
2. Developed an evaluation framework for the research portfolio;
3. Conducted a round of regional workshops to explain their new evaluation frameworks to research and extension staff;
4. Contracted an evaluation firm to conduct training for a selection of researchers who had submitted preliminary research proposals for the 2004-05 year; and
5. Through this training the researchers developed an evaluation plan which was included with their full project proposal submissions.

The GRDC have introduced an outcomes-oriented project development system. Researchers have been heard to comment that the outcome based project application format introduced by GRDC increases the complexity of applications but that it does help them to think about the larger context and impact of the research they are planning.

The GRDC funded northern Farming Systems projects used a shared evaluation approach. Project leaders believe this contributed to developing the team research approach and focus and aided researchers in learning about the farming systems research process that they were using (Lawrence 2004 *pers. comm.).* The leaders of these projects believed that evaluation and reflection were critical to participatory RD&E. However, team members rated evaluation as one of the least important aspects of the project (Lawrence, Carberry *et al.* 2001). These authors felt that team members had little previous experience with evaluation and saw little incentive for evaluating their own activities.

CAPACITY

Culture will develop the interest, belief and commitment for evaluation but not necessarily the skills required. Evaluation capacity building is also needed to develop skills and confidence in the use of evaluation frameworks, the conceptual approach to and benefit of evaluation, the nature of social research and inquiry and the methods used to gather and interpret evaluation data. Training and project support was a clear part of the success of the evaluation capability program in DPI Victoria. However, McDonald, Rogers, et al (2003) stress the importance of balancing both the supply and demand for evaluation capacity building.

The University Cooperative Extension services in the USA invest considerably in developing training programs, resources, web sites and personal support to enhance the evaluation skills of extension agents. Several university agricultural extension programs have an evaluation unit employing a number of evaluation academics and practitioners. Excellent resources for evaluation are available through the websites such as the Pennsylvania State University (http://www.extension.psu.edu/evaluation/) and the University of Wisconsin Cooperative Extension Program Development and Evaluation (http://www.uwex.edu/ces/pdande/).

Considerable variation in the frequency of use of evaluation techniques was found between researchers in the sugar industry, with the most common technique being checking of progress against project milestones (Henderson 2003). 78% of respondents to this study felt it would be valuable to learn how to use evaluation tools to improve the efficiency and effectiveness of collaborative research.

Who Evaluates?

Accepting that research agencies are now required to have an evaluation approach, Research and Development Corporations and CRCs have essentially these options:

- Include evaluation reporting as a requirement in the project application and reporting process – this puts the responsibility on the project leaders and their agencies. Program level evaluation would need to be collated from the project evaluations, perhaps by staff of the RDC or CRC.
- Contract evaluation studies to external evaluation firms;
- Build internal capacity by employing an evaluator on staff, funding an evaluation project or training existing staff and projects.
- A mix of the above.

Rather than increase the budget for external evaluations, most public sector agencies have focussed on making changes within the agency by developing staff skills and sometimes broader organisational change (McDonald, Rogers *et al.* 2003). This internal approach is sometimes done because of a perception that there are too few external evaluators available (Leeuw, Toulemonde *et al.* 1999) or to save funds. Many organisations invest in developing internal evaluation capability in the aim of broader, long term benefits which have been demonstrated by (Fetterman 1996) and others.

Relying solely on self-evaluation by programme staff and managers has a high risk of servicing the interests of those stakeholders only (Scriven 1996). For this reason, McDonald *et al* (2003) argue that internal evaluation capability should include internal approaches together with external evaluations and on-going monitoring. A part of the internal capability required will be in managing and using external evaluations.

This latter point is also identified by AACM International (1997) though in this case they do not advise any other elements of internal evaluation capability. Their recommendations for the then Land and Water Resources Research and Development Corporation (LWRRDC) suggest levels of involvement with evaluation:

- The Board, who are responsible for allocation of investment be a part of *ex ante* evaluation and be aware of the results of progress evaluation and monitoring and any *ex post* evaluations.
- LWRRDC staff ensure that the implementation of the evaluation meets the defined standards and levels of accountability.
- LWRRDC program managers should understand the roles of evaluation, the application of specific evaluation methods and the integration of these into management of programs. They should be skilled to manage contract for external provision of evaluation services.
- Evaluator should be independent from those with responsibility for managing the program.

Interestingly, this view does not include the researchers or other stakeholders in the process. Perhaps the perspective is restricted by the author's own paradigm which strongly relies on contracting external evaluators. It could be argued that this model is perhaps appropriate where the primary purpose is accountability and resource allocation but perhaps not so well suited to project planning, improvement and learning purposes. This may have matched well with the nature of LWRRDCs funding partnerships.

In practice, a mixture of internal and external approaches is often used, particularly where a funding agency is interested in building the capacity of the researchers with which they regularly partner. Who does the evaluation often varies depending on the level in the output-outcome-impact chain. For example, Table 6 depicts evaluations of extension activities related to Integrated Pest Management (IPM) in the cotton industry. Higher level evaluations have been done by external evaluators, evaluation of attitudes were done by extension staff under the guidance of external evaluators and activity level evaluations were done solely by extension staff.

Table 6 Evaluation of extension activities at multiple levels related to IPM in the Cotton Industry mapped against Bennett's Hierarchy. Adapted from (Christiansen, Pyke *et al.* 2003). Those marked with shading indicate evaluations conducted externally or through collaboration with an external evaluator. Other evaluations were conducted by a staff member or group of staff.

Bennett's level	Evaluations conducted				
Outcomes - Impacts	- Analysis of total pesticide usage data - Water quality data - Record of social complaints - Benefit-Cost Analysis				
Practices	- Benchmark Survey	Attitudes to IPM - Focus groups 1997 & 2001			Regional Tipping Out Trial Impact Survey
Knowledge, Attitudes, Skills & Aspirations	- IPM Short Course follow up interviews		AWM groups phone survey	Insecticide Usage Comparative Analysis Evaluation	
Reactions	- IPM Short Course pilot - Cotton Tales Surveys - Information Resources Survey - Field day evaluations				
Participation	- IPM Short Course Numbers - Numbers of AWM groups - Cotton CRC website hits				
Activities	- Project reports				
Inputs & Resources	- Annual reports - Project Budgets				

Clear communication of the purpose, goals, responsibilities and expectations of all stakeholders will aid a participatory evaluation process that involves multiple stakeholders (Lawrence, Christodoulou *et al.* 2003).

GRDC has developed in-house capacity for the economic analysis and allocative modelling of investments that works with their outputs and outcomes framework for projects (GRDC 2002).

Evaluation Use in RD&E

There is little published evidence of how evaluation has been used in RD&E in Australian agriculture. Many agencies and R&D groups are at the early stages of developing evaluation. For RDCs this appears to have been largely driven by the need for accountability to the Commonwealth. Evaluation and reviews are also being used to plan research investments. CRDC have indicated a desire to increase researchers' understanding of the contribution of their research towards outcomes and impacts (Pyke 2004 *pers. comm*).

In reference to Table 6 , some of the value of these evaluations was in providing feedback to individuals to aid their professional development and to refine their activities. More specifically, the recommendations of the IPM focus groups in 1997 led to the development of an IPM Short Course for cotton growers and a set of IPM guidelines. The 2001 focus groups endorsed the input of extension staff in supporting Area Wide Management groups as a core part of their IPM extension role. The outcome and impact levels have been reported in a recent CRDC annual report (CRDC 2003a).

RESEARCH METHODS AND CONTEXT

This research builds on approaches previously initiated to build evaluation capability within the cotton extension network. Previous evaluation activities with the extension network have included training, frameworks, evaluation exercises and a skills survey. The need for evaluation had been identified by and regularly communicated to the extension network by the CRDC board. This study has specifically focussed on the research sector as it has not been a part of this previous activity, but will equally need to be able to demonstrate value from research investments.

A number of contextual issues surround this study. A Benefit Cost Analysis had recently been conducted of the Cotton CRC. This analysis was undertaken by an independent evaluator who met with several research and extension staff. CRDC had also recently undertaken reviews of selected research programs. CRDC and the Cotton CRC have been under increasing pressure from the Commonwealth to demonstrate the value of research investments. At the time of the study many Cotton CRC researchers were involved in the development of a proposal for a new CRC. The guidelines for new CRCs, that they were all working towards and should have been well aware of, required a significant focus on clear economic benefits for Australia.

In addition, my own role in the Cotton CRC had recently changed from extension coordination to one of providing evaluation and extension support and research for both research and extension programs. Most researchers were not well aware of this and it was hoped that this study would help to raise awareness amongst researchers as well as set priorities for this changed role.

The research plan evolved through exploring the emerging needs for evaluation and triple bottom line reporting of both extension and research. In developing evaluation skills in the extension network, the extension group recognised the value of evaluation frames as planning tools. With this in mind, it was hoped that an effective evaluation strategy that involved both research and extension could enhance the linkages between these sectors.

In an initial attempt to explore this potential, a researcher with an interest in evaluation worked with me to try to map out achievements against Bennett's Hierarchy. We attempted to do this for each of CRDC's planned impacts (CRDC 2003b) by including the contributions and impacts from research and extension projects. In attempting this, we found it difficult to relate some of the research outputs directly to Bennett's seven levels. From this we developed a "continuum" from: contribution to knowledge through to sharing and integrating knowledge; developing ownership, confidence and understanding in the industry and to practice change and impacts (included in Appendix 2). In looking at this we identified that some projects covered all aspects but most focussed on a particular area. It was potentially a way to identify where projects fit and potential gaps in linking basic research to adoption.

This conceptual framework was discussed with a few researchers, in Toowoomba and Narrabri, who identified well with it. In particular, researchers liked the concept of a continuum which identified team roles. That is, whilst all industry research may be expected to ultimately lead towards impact, each of the steps in achieving that need not be done by each project. The framework could however clearly identify where linkages were needed. On this basis, this project initially aimed to develop and test this conceptual framework by mapping existing data against the framework.

Taking this concept further, an initial research plan was developed (Appendix 1) and circulated for comment to leaders of the Cotton CRC extension program, CEO of the Cotton CRC and a CRDC research program manager. Feedback on this suggested that, whilst such a framework could be very useful, it would be helpful to first identify more broadly amongst researchers the need for such a framework and researchers' current approaches to evaluation. This was done through interviews of researchers parallel to a review of available data. These differing approaches have provided some triangulation to increase the confidence in the findings of this research.

Review of existing data

Copies were obtained from CRDC of the 2002-03 annual research reports submitted to CRDC by all funded researchers. (This excluded plant breeding and biotechnology research reports, which are confidential.) The annual reports were briefly reviewed to identify the potential for gathering data about research outcomes and impacts from these reports. This was done with a view towards mapping against an evaluation framework the outputs, outcomes and impacts relating to a few specific research program areas, such as IPM, Diseases or Water Management.

Interviews of Researchers

To gain insight to researchers' perceptions about and practice of evaluation, twenty-four researchers were interviewed. This represents almost half of the Cotton CRCs total number of researchers. The interviews consulted a cross section of cotton researchers from four locations: Toowoomba, Armidale, Narrabri and Sydney. Most, but not all, of these were participants of the Australian Cotton Cooperative Research Centre. They were from five organisations: Queensland Department of Primary Industries and Fisheries (DPI&F), NSW Agriculture, CSIRO (Divisions of Plant Industry and Entomology), University of New England (UNE) and The University of Sydney (USyd). The Cotton CRC also involves small numbers of researchers in Central Queensland, Northern Territory, Western Australia and Victoria who were not able to be consulted in this study.

The semi-structured interviews, each of about an hour in duration, were conducted at each of the researchers' offices between March and May 2004. Interview questions (Appendix 2) were developed with input from research managers and with advice from social researchers. To make the most of the interview schedules, questions relating to linkages between extension and research were included. These provide some general context for this study but will otherwise be reported as a separate study. In light of time constraints of some interviewees some interviews were condensed to a subset of these questions.

Transcripts of the interviews were collated for subsets of researchers, grouped by location and organisation: Toowoomba (DPI & F); Sydney (USyd); Narrabri (CSIRO); Narrabri (NSW Agriculture) and Armidale (UNE). These collated interview transcripts (each of 3-7 researchers) were emailed to the relevant researchers asking for feedback about any perspectives that were not adequately or correctly captured. The interviews were subsequently analysed by emerging themes and as clusters as detailed in the next section.

FINDINGS

The findings from the study are presented in light of the three key elements of the study. That is: the perspectives and interests of research managers; an overview of the usefulness of the annual reports for mapping project influence; and the key themes that emerged from analysis of the interviews.

Research Managers' perspectives

Cotton research is largely funded and managed by the Cotton CRC and the CRDC. The underlying evaluation questions for these organisations were: "Are we getting a return on investment? What can we improve? What are the project achievements and outcomes and what can be done better along the way?"

Previous work has outlined (Christiansen, Pyke *et al.* 2003) key considerations in choosing a framework for evaluation of cotton extension programs:

- Useful for extension evaluation and extension planning.
- Flexible for use with a wide range of issues and evaluation methods.
- Simple to use and to communicate.
- Aligned to the key outcomes of our funding bodies.
- Helps extension staff to review, learn and improve.
- Encourages staff to be mindful of outcomes and impacts when planning activities.
- Stimulate creative approaches to evaluation.
- Not too prescriptive as this may be seen as a chore and may not suit some evaluation methods or personal preferences.
- Efficient use of evaluation resources.

Further discussions with senior managers of both of the CRDC and Cotton CRC highlighted key elements they would like from research evaluation.

A CRDC perspective:

- A framework or process like the Logic Model that will help to get researchers thinking through to impacts.
- Feedback about impacts in industry that can be attributed to research.

- Encouraging researchers to think about what could happen in the industry if the research was applied, and how this might link through to extension plans.
- What would change for a grower if that bit of knowledge from a research project was applied?
- Through research and extension synergy, be able to identify and communicate – What are the key messages from the research? What's in it for the industry?
- If researchers don't do the evaluation, we'd like them to at least be thinking about what the impacts might be. [This comment evolved from discussion following initial interviews.]

Cotton CRC perspective:

- To develop a process to demonstrate what the projects and programs have achieved in terms of outputs, outcomes and impacts and how could this be demonstrated and monitored (even if it is not done by the researchers*).
- Approach evaluation not as a fault finding exercise, rather as a means for continual improvement, a way to demonstrate progress and success and as a means for researchers* to justify why projects should be refunded.
- We need to improve our internal evaluation, but will still need independent review and assessment, especially to demonstrate an economic return on investment. It all gets down to dollars, even though for many things this is difficult.
- Understand how to measure triple bottom line outcomes and impacts. Each specific project does not have to necessarily address all three, but the Cotton CRC as an organisation does.
- It will often be better to evaluate outcomes and impacts of a group of projects, both research and extension.
- Surveys maybe an overused form of evaluation in our system.
- We need a couple of simple frameworks.

* Note here that the word "researcher" referred to researchers, extension officers and educators.

In progressing these needs, it was considered helpful for this study to look at:

- Researchers' perceptions of outcome and outputs.

- What do researchers believe that their projects have achieved?
- Short term and long term impacts.
- What are the differences in approach to evaluation and continuous improvement across organisations, locations and disciplines?
- Developing a process for the Cotton CRC to evaluate outcomes.

These discussions contributed towards the development of the key research questions and the research context.

Review of existing data

Annual project reports submitted by researchers to CRDC in 2003 were briefly reviewed. This identified that these reports provide information about outputs but very little if any indication of adoption or impacts. In response to a question relating to the impacts of the research, researchers generally provide a statement of the expected or probable impact of the research.

It would not be feasible to map outputs, outcomes and impacts of research solely from the data contained in these reports. However, the output level information and the statements of the expected impacts in industry can be used to guide evaluation of higher level outcomes and impacts.

Key Themes Emerging from Interviews

The issues emerging from the interviews are identified and discussed in relation to key themes. The earlier themes identify the current situation whilst the latter ones describe opportunities for moving forward.

The concepts surrounding what researchers perceive as the role of their research and the values that underpin their research practice provide an insight into their attitudes towards evaluation. These values will need to be considered in planning an evaluation strategy that these researchers can relate to and contribute towards.

THE ROLE OF RESEARCH

The researchers interviewed held varied perceptions of the role that their research plays. These views fit a spectrum of the variable depth of research embeddedness in the change processes. At the one extreme, research was positioned as an information provider whilst at the other research was seen to be integral to facilitating change in industry practice. For some researchers, contributing to new knowledge went hand in hand with changing practices, whilst for others the place of research was to provide information and not to 'enforce change'.

These core roles of research can be described in more detail as:

1. *Information provider* This role focussed on providing accurate, valuable information; identifying knowledge gaps and filling those; regional information for decision making; and providing information to extension to further industry understanding. The information provider role was not restricted to industry stakeholders. Several researchers identifying it as important for research to contribute towards the world knowledge base, to advance scientific knowledge and to pioneer understanding. A range of researchers aligned themselves with this information provider role, particularly those at Sydney University.
2. *Integration of knowledge* A key part of research, considered by a small number of researchers, was in the integration of understanding different components of the farming system. They display a problem oriented approach of seeking solutions to address problems and issues in the industry, sometimes at a regional basis. Particularly those researchers involved with disease management, farming systems and fibre quality, noted this role.
3. *Developing tools and management options* Tools and management options were considered to be a way to deliver science to industry in practical ways and to ensure that science can aid decision-making. Research had a role both in providing and assessing management options and in independently evaluating the fit in the system of options developed by others.
4. *Relevance to Industry and Guiding Investment* This role, identified explicitly by only a few researchers, included guiding the overall relevance of research

investments, and ensuring that the research was conducted and communicated in ways that made it relevant to the industry members. Some research informed research investors of areas of priority needs for the industry currently and in the future. This included testing possible frameworks for natural resource management allocation, reviewing research knowledge to guide future research needs and in focussing the research being undertaken and the way in which it is communicated.

5. ***Building Capability*** This role for research was perceived by the education sector who aligned with a key role in building the capability of people in the industry, through building both a learning culture and industry skills. Other elements, related to maintaining and building research capability for the industry included: 1) for the research to be rewarding and enjoyable for the researcher; and 2) strategies to maintain staff and capacity and to build research skills within the context of short term contracts.
6. ***Improving Industry Practice*** A number of researchers identified achieving change in industry practice as a key role of research. This ranged from the adoption of specific practices, through to objectives such as a reduction in chemical usage and contributing towards the overall sustainability of the industry, which in turn had flow on impacts on the sustainability of the rural communities. It was the IPM researchers who most regularly linked the research role to practice change.

Other perceived roles of research included moderating risk through providing independent science and promoting to community the good practices and science adopted by the industry.

RESEARCH VALUES AND MOTIVATION

Some of the 'labels' that stand out from the discussion of research values include: Honesty, Integrity, Rigour, Measurable, Passion, Usability, Change, Sustainability, Belief, Visibility and Learning.

Quality science that adds to knowledge was a core value underpinning the researchers' approach. This included honesty and integrity in data, technical merit and the application of basic principles from that field of science. There was concern expressed about an emerging trend to take shortcuts in research, lacking respect for the scientific methods of adequate replication, analysis, design and rigour. Reading widely about the issue and collaborating with researchers elsewhere (outside of cotton, both nationally and internationally) was considered important for quality science.

For many it was important to publish the science for multiple audiences, as this indicated that it is sound to peer, international and industry review. This was particularly true for the university and CSIRO researchers. Some commented that as well as helping them to understand the information better, publishing through a peer reviewed process gave them confidence in the soundness of the science. For this reason, a researcher expressed reluctance to release information to industry before it was scientifically published, a process that may take up to two years.

Honesty and belief in the science was a value that was further expanded by two researchers, both of whom had experienced difficulty in communicating scientific results that were not widely accepted or desired. They stressed the importance of reporting on the science, not on what people want to hear. Going into the research with an open mind, not blinkered by believing all of the 'existing myths and truisms', had been shown to identify critical issues for the industry. At times this had been at some personal cost for those researchers.

A passion for the work and enjoying doing it was identified most commonly by the more senior researchers. More broadly, a passion, belief in the goal and personal ideology about the research was evident. This included personal commitments to the field of work, such as IPM, or more broadly to the need for sustainable production of food and fibre. This was coupled with a belief in the positive contribution that science could make to these fields. Some mentioned that they felt privileged and lucky to be able to work on contributing towards something that they personally felt strongly about.

"I'm personally committed to IPM – perhaps biased towards trying to convince farmers that it is a good approach. The whole ethos of my work is about the adoption of IPM."

Another expressed the interest in working with the industry rather than against it as a means of achieving positive change for nature conservation.

In addition, some underpinned their research with the goal of 'doing something that directly helps growers', for the science to be useful. This included querying whether and how the proposed research would deliver benefits and real change to the industry. Ensuring farmer profitability, which in turn led to benefits for rural communities and Australia at large was another value. These researchers also felt that they tended to be extension oriented.

Visibility in the industry was considered to pay dividends for researchers in terms of continued funding and personal appreciation through interaction with industry. It was considered to be important to be seen to be out working in the industry. Whilst tended to lead to becoming a "jack-of-all-trades" in working in the industry, this researcher also commented that this approach does not readily lead towards scientific publication.

Fundamental for the educators was the need to provide a good learning experience for students. Achieving this required balancing the students learning needs with the industry needs from the student's research.

Motivation and Conflicts

"I <u>love</u> it!"

Enjoying the work, a quest for the science and discovery, fun seeing results for industry were key motivators for researchers. Many commented that they were driven by a desire to figure out how things worked, by science and knowledge itself, by the 'thrill of the chase' and by striving to do the best that you can.

Another strong motivator was the positive feeling about 'making a difference'. Reinforcing the core values, several were motivated by seeing change in the industry, providing cutting edge research to make the industry more competitive, by doing rigorous research that is also applicable to growers and by reducing the stigma that is associated with the broader public's perception of the industry's environmental management. Direct thanks from growers and industry awards were appreciated.

Some experienced various conflicts between their personal motivations and their organisation's expectations. This was particularly true for those scientists in CSIRO or NSW Agriculture whose focus lean towards facilitating change in the industry. These conflicts most commonly were associated with a pressure to publish scientific papers in order to achieve salary progression. It was commented that much industry-applied research, which may be significant to the industry, is often difficult to publish in scientific journals and that their organisations gave little credit for the time invested in communication with industry or in the changes that have occurred as a result of the research.

There was perceived to be disparity between the priorities for research funding and organisational rewards. On the one hand, achieving on-going funding (which for many was essential for on-going employment tenure) was influenced by the scientist's communication with industry and contribution towards changes in practice. On the other, salary incentives for researchers focused on scientific publications. Some researchers had appreciated positive feedback from industry about the great research work being done, whilst also experiencing negativity from managers due to low numbers of publications. A few more experienced scientists commented that one of the keys skills in industry based research is to balance these sometimes conflicting demands. There was some request from newer staff for some training or mentoring of how to best balance these needs.

Note that discussion about this point with CRDC revealed that CRDC also values scientific publications and would like to see more peer reviewed publications from some researchers.

Some scientists found it challenging to balance their roles as managers with their ability to do research. Juggling short term contracts posed a challenge for some researchers, both for their own projects and tenure and for managing and maintaining the staff in their teams. It is interesting to note that only a few of the many researchers on contract positions viewed this as a difficulty, this was likely influenced by specific circumstances an experiences. There was also comment of a lack of appreciation for the time needed to seek research funds from other sources.

It is interesting to note that, for the most part, university and DPI&F staff experienced little if any conflict with their organisation. For example, the 'pressure to publish' was not regarded as such a conflict – in DPI&F because it was less essential and in the universities because the researchers did already focus on regular publishing. Provided that they brought in research funds, each of these groups felt they were largely left with a free reign to conduct their projects. Some indicated that this was probably because they had a track record of delivering and being successful. Some DPI&F staff commented that their organisation's goals matched closely with those of the industry research so it was easy to 'package the research' to fit DPI&F needs without compromise. Another commented about 'not really knowing what the organisation expected', this tended to hold true for the two state agencies and the universities. In the university context, this was in no way seen as a problem as the very nature of academic research positions was the freedom to explore a scientific area of interest, with the only requirement from the university being to attract funds, to publish, to deliver quality teaching and to provide service to industry, the profession or the community.

One concerning perspective was that a desire to 'do the best I can' sometimes conflicted with a creeping attitude of 'doing just what's required.

Pure and Applied Research

Overwhelmingly the researchers interviewed indicated that their work was solely or predominantly applied research. Some indicated that they had a mixture of some pure and some applied, though in most of the pure cases it was to contribute to applied research in some way. For example, pure research may be undertaken to develop a diagnostic tool that would then enable applied research to be done. Some have followed a continuum within a project, initiating with pure research and then following this through to very applied field application and development of management strategies.

Interestingly, the perceptions of what is "applied" and what is "pure" research seemed to be quite varied across the researchers. For example, some felt that "applied" is any research that is conducted in an industry context, or has the potential to be used by industry in some way, perhaps through further research. Others regarded "applied" as being more about application in the field, meeting industry needs and informing or facilitating change.

Researchers generally felt that CRDC seemed to focus predominantly on applied research. This was linked largely to the influence of ACGRA who are after answers to problems. Many commented that CRDC would invest in pure research where they could see the potential value of the knowledge. A comment was made that CRDC seems to perceive the state agencies as providers of applied research, and CSIRO as a provider of some more basic research. Some university researchers noted that it is almost impossible to gain funding for pure research through the RDCs other than through student scholarships and that it isn't possible to access Australian Research Council grants for this type of agricultural research as it is regarded as too close to industry.

A few commented that the Cotton CRC had more scope for "blue sky" research. This was attributed to the fact that the Cotton CRC itself will be evaluated on the quality of its science and so it needs some fundamental science. Some of the longer term Cotton CRC investment in what was initially relatively pure research is now being seen to deliver benefits.

Nature of Research Practice

The majority of researchers felt that their work was team based or a mixture of team and individual approaches. Some felt their research was more individual, with collaboration with others at various stages of the project life. The association with a sense of team ranged very strong to uncertain. Some teamwork was seen to occur in an ad hoc, unplanned manner. Some teams involved the whole industry, whilst others focussed on research. Several noted that they were involved with several teams at industry, organisation, national and international levels. There was also a practice of linking with necessary skills to form teams as required.

Across the board researchers approach their work in a programmatic manner. They all felt that their work builds on previous research and provides building blocks for future research – be it their own or that of others. Some commented that funding arrangements dictate a project based approach, so they have operated as a suite of projects in the broader agenda. There was comment of a lack of a long term strategic plan of direction for research in IPM, with limited discussion at the end of project terms of achievements and future needs.

Elements of Success

The most common factor constituting a successful research project was:

"One that is completed and meets most or all of its objectives."

This included the project meeting its original goals and reaching useful milestones. Some added a problem solving context of solving the question, providing information and knowledge about a particular problem.

Other elements of success from a research project were identified as:

1. *Scientific Success* Published scientific papers, acceptance by peers.
2. *Industry Success* Adoption by industry at a high level / 100%, somehow has impacted on cotton production, relevance to industry, outcomes that growers are willing to adopt, people using research and building on that finding.' This included a concept of research that 'enables' through providing management options that are applicable, they work, they fit with possible practice and they can be adopted economically. *"See or hear of growers taking action based on the research that they felt they had decided themselves."* Appreciation from the grower and research put into a forum where it can have an impact were elements of successful research, as were the adoption and performance of tools, practices or varieties where intended.
3. *Research Success* On-going funding support for a future project.
4. *Personal Success* Wanting to continue in research, a feeling of satisfaction with a 'job well done' and building own perspectives and confidence in the area were clearly identified by a few researchers.

STAKEHOLDERS OF RESEARCH

The majority of researchers identified growers and consultants as key stakeholders for their research. Several focussed specifically on the growers and consultants who they collaborated with in research trials. Many also identified as stakeholders the research investors: CRDC, GRDC and Cotton CRC. Several CSIRO, DPI&F and NSW Agriculture staff identified their organisations as stakeholders but interestingly no university researchers did. Some researchers also identified stakeholders in the 'people of Australia' (public benefit), Catchment Management Authorities, Chemical Companies (mainly for contract research work), the Scientific Community, other Researchers, Extension and Students. One researcher also identified other secondary stakeholders as being everyone along the cotton value chain who would be impacted by the productivity of cotton , such as cotton ginners, transporters, regional communities and the national economy.

These stakeholders were seen to access research through direct interaction with researchers, articles in the Cottongrower magazine, farm walks and field days, through the extension network, at the cotton conference, CCA forums and other seminars,

Messages were seen to get out quickly as trials were conducted on commercial farms and the grower and consultant collaborators usually interacted with the research. Some commented that the cotton industry is very rapid to pick up new ideas from research and try them.

Feedback

Many researchers readily received feedback about their research from discussion with growers and consultants. There were numerous ways by which people 'tap into the industry' including forums, conferences, field days and one-on-one contact whilst doing other things. Many researchers, particularly those in Toowoomba and Narrabri received frequent phone calls from growers and consultants. The growers and consultants involved with the on-farm trials were particularly important sources of feedback. Some actively sought this feedback whilst others didn't.

> *"People talk – farmers are not reticent in letting you know what they think."*

A few scientists considered feedback from growers and consultants who had tested a recommended practice to be an important part of the research process. Feedback from ACGRA and CRDC was also important, partly because it represented a broader set of views, but was considered by some to be less forthcoming than that direct from growers and consultants. Some sought feedback from Cotton Seed Distributors (CSD). Very few researchers used independent evaluation as a source of feedback.

Many identified a potential for the extension network in playing a stronger role in facilitating the two-way flow of information, feedback and priorities between research and industry. Most researchers felt that these linkages between research and extension could be improved.

Feedback about the science was gained through peer reviews, generally of publications, through CRC Reviews, and for some CSIRO researchers from discussions within their group. One researcher expressed concern about the lack of a process or a culture for frank, critical appraisal of research. It was felt that the quality and relevance of the science would benefit from more constructive debate about the scientific approach used, the usefulness of the research and synthesis with other research. This included a need for frank feedback from both other scientists and the RDCs.

In considering how to react to feedback, it needed to be first evaluated and considered. Particularly constructive comment could provide new sets of ideas and ways of thinking about an issue. Feedback was also used to develop priorities, help contribute to and understand the research, adapt programs and modify research if it was off-track.

> *"Feedback from industry and colleagues helps to shape my thinking and approaches. I'll gather it and follow it up with industry leaders."*

CONCEPTS OF ECONOMIC, ENVIRONMENTAL AND SOCIAL IMPACTS

Most researchers had thought to some degree about the triple bottom line impacts, and generally hoped that it would impact on all three in some way. A general concept of aiming towards sustainability across each of these was regularly expressed. There was little detailed knowledge of the impacts of the research on each of these bottom lines, and some commented that they needed to be able to understand this better. People had most difficulty in understanding what the social impacts may be. Table 7 presents the various thoughts and ideologies about the impacts of research on these three pillars of sustainability.

As observed by one researcher, many researchers are not comfortable with considering environmental, social and economic factors.

> *"... environmental, social and economic elements pushes people out of their comfort zones."*

Table 7 Researchers' expectations of the triple bottom line impacts of research.

Economic	Environmental	Social
Increase yields and gross margins	Reduce pesticide usage Ecologically sustainable practice	Involved with AWM groups Research empowers people – pride amongst collaborators Community ownership
Savings on cost of control – may be offset by increased management cost eg trap crops	Expect reduced insecticides, especially broad spectrum has benefits on & off farm. Cumulatively could be meaningful across a region -> water quality, biodiversity	AWM impact on concept of a community issue. Benefits of growers and consultants getting together and talking – networks may be used for other things as well
Systems thinking		
Profitability Industry competitiveness	Care for the environment	
	Anecdotal stories of more beneficial insects and birds	Confidence building in interpreting risk of IPM
Reduce risk of yield loss or delayed maturity Increase profitability	Resistance management Less insecticide use Reduced risk of off-farm pesticide movement	?? Ensuring profitability -> thriving community Public perception
Save $ with fewer sprays		Community perception – keep industry from being shut down due to use of hard chemicals
Native vegetation and beneficial insects – potential impacts.		
Longer term, diffuse through research investment		Biodiversity review demonstrates to community industry action & leadership.
Analysis of irrigation systems		
Retained production (estimated $70 mil loss across Downs previously) – flow on effect.	Not using chemicals to manage the problem	Restored confidence in cotton production. Employment
Align with Cost-benefit – what would have happened if research wasn't done?		Provides answers about criticism of the industry.
Improve yield by reduced waterlogging and improved WUE from moving storages	Reduce waterlogging and salinity	Confidence, reinforcing growers ideas about the problems.
More cotton with less water	Better returns per megalitre	Better perception about the industry as aiming to be more efficient
Not so much	Hopefully environmental sustainability	Unsure
Have independent BCA numbers on this.	Improved WUE with varieties Reduced need for insecticides	Multiplier based on industry expenditure
Setting up systems for environmental benefit that don't cost more.		
Hopefully more sustainable across all three		

MEASURING AND REPORTING OUTPUTS, OUTCOMES AND IMPACTS

Measuring

As a rule, researchers don't measure the adoption or impacts of their research. Some gather a general feeling about change in attitudes and practices based on discussion and interaction with industry. Some commented that there are some subtle changes in the attitudes or actions of the growers and consultants they interact with whilst other changes can be quite dramatic. Several said that they wouldn't know the level of adoption of their tools or recommendations, though generally they believed it had an impact. Some commented that there was rapid uptake of their research. A large number queried whether it was in fact possible to measure this at all, and many suggested a survey would be needed for this.

Some researchers, mainly in NSW Agriculture, had worked with economists to assess the economic benefits of their recommendations. They commented on the need to appoint another economist to the Cotton CRC (a position left vacant when the previous economist left).

A few researchers commented that they referred to the studies of attitude change and practices that have been done by the extension network and the reports from the Cotton Consultants Australia (CCA) surveys. These researchers were mainly working in the IPM field, an area that has had more evaluation data gathered and studies available. One researcher commented that whilst they gathered qualitative data, he wouldn't think that they had measured adoption and impact as they have no quantitative measures.

Research programs that link with commercial partners receive measures of the use of their products (cotton varieties or semio-chemical) from that commercial partner. These were also identified as relatively simple factors to measure as they are largely based on sales and hectares planted. Provided that varieties performed as expected, impacts could be calculated from the area planted.

Other than these who link with commercial partners, only two research programs confirmed that they attempt to measure adoption. In one case this was done by an external evaluator. In the other, the research team themselves tried to measure to some extent practice change and the changes in knowledge and understanding that underpinned that practice change. This latter was considered to sometimes be essential, as the actual change in practice may not be seen within the timescale of the research project/s. It is interesting to note that one of these researchers had previously worked in a department that had a strong evaluation strategy for research. Another researcher commented that adoption and impact is measured directly where possible but that much research is not clear and measurable.

Reporting

It was noted that reporting on outputs, outcomes and impacts is increasingly being required of researchers. Some only reported on these if they were required to. Researchers generally were readily able to report on outputs, a few reported outcomes and most didn't report any impacts unless they were obvious.

There was variability in the interpretation of what is meant by outputs, outcomes and impacts. This confusion is consistent with similar findings from a survey of the cotton extension network (Christiansen 2003, unpublished survey report). For some researchers, and for the university systems, impacts are regarded as citations, student evaluations and research community involvement. Some regarded publications themselves as an impact whilst other viewed them as an output. Table 8 lists some of the classes of outputs, outcomes and impacts that researchers identified with.

Table 8 Classes of Outputs, Outcomes and Impacts reported by researchers.

Outputs	Outcomes	Impacts
Products	Tangible & Intangible *eg soil information, interpretations, maps*	Not reported
Publication, articles, changes to website	Final report - summary of what research was and what a farmer benefited from it (case studies).	Not really reported
Yield, cost of production, new tools, beneficial insect activity	Yield, gross margins, financial	
		For university – citations, student evaluations, research community involvement

The reporting process most regularly mentioned by researchers was the interim, annual and final reporting of projects to the CRDC and Cotton CRC. Some also reported to other CRCs or GRDC and a few to their organisations. These are used to report against milestones and objectives. It is interesting to note the differences in perceptions of what is expected of these reports. For example, some researchers commented that they did not report on outcomes and impacts as there was little requirement to. Conversely, another researcher commented that they do attempt to report on impacts and outcomes as this is the 'expectation of the RDCs':

> *"We try to report on outputs, outcomes and impacts as guided by stakeholder's reporting requirements. We are expected to include these in final reports….I expect that everyone does some form of reporting on these, though due to the lack of clarity of what is expected it generally tends to be just a stab at what the impacts might be."*

A few had noted that the most recent CRDC annual report format (February 2004) had included a question about the impact of the research. Lacking data, researchers answered this question with a guess, a rough estimate or an expectation.

For some, reporting outcomes and impacts was driven not so much by a requirement as by a commitment towards demonstrating to their funders that their research is making a difference. Many identified multiple reporting mechanisms, adding more interactive reporting such as: at the annual Cotton CRC Review, at Cotton Conferences and through working groups such as FusCom.

The lack of clarity and understanding of *how* to report on outcomes and impacts was a recurring theme. Researchers identified the difficulties of reporting on these 'in a way that makes sense' and in generating good quantitative data about research success. This was coupled with an identified lack of skills in measuring attitudes and change:

> *"I don't want to have to report on outcomes and impacts... How would you really do it? How can you measure that?"*

It was also noted that it is more difficult to identify impacts and outcomes of that research that does not lead towards a clear 'product'. Some mentioned that, due to the nature of their research they are only now having outcomes to report as the early work was exploratory and developmental.

Report Format and Use

The current formats for CRDC and Cotton CRC reporting were generally considered to be suitable for research programs though less well suited to other work such as education. Some felt they could be improved, some liked them and several commented that they wouldn't like to see them become too complex. The flexibility in the current reporting format was appreciated.

However, several researchers expressed concern about the lack of apparent use of the reports provided to CRDC. Researchers commented that a lot of information is contained in a final report, representing a considerable investment of time. Some of the specific concerns included:

- This information from research programs may be being lost by the reports becoming 'buried'. Can they be archived in an accessible format?
- Some reports include important implications for priority setting which may be being overlooked, or viewed too late.
- If the reports are not adequately used, researchers will stop putting effort into doing them well.
- After including evaluation results in a final report, one group of researchers sought feedback from the RDCs as to whether this had been valuable and suitable. They found it difficult to receive this feedback from the RDCs.

- Some researchers felt that they can't rely on the reports as a feedback mechanism.
- The only feedback received is a standard form letter to say that the report has been accepted.
- There is no mechanism for regular or spontaneous feedback from RDCs and little feedback is received unless you know the research managers well.

The degree of concern about the usage of reports ranged from a mild curiosity about whether the reports were really being used and how, to a feeling that the CRDC reports are wasted because they are not being read and so the reporting is not a helpful process. To a large degree, this frustration appeared to be based on a view that the reports could be very useful and a desire for frank feedback from research investors. There was considered to be significant opportunity for improvement by increasing the feedback on, accessibility to and use of these reports.

Many researchers were keen to receive more feedback from the funders about their reports and their research in general. Several considered it would be helpful to receive more feedback on the reports submitted. One researcher commented that the CRDC reports are quite helpful as they are also sent to ACGRA members. However, he was concerned that, due to the high number of reports each person received, they wouldn't be able to do more than skim the reports and perhaps look in more detail at only those of personal interest. There was also suggestion that these research reports could be reviewed to identify data for evaluation reporting.

Evaluation frameworks

As far as the researchers were aware, none of their organisations utilised frameworks for evaluation. Very few researchers were aware of what these frameworks may be and how they may be used. Several researchers referred to the CRDC, GRDC and Cotton CRC reporting formats. Universities have a system for researchers to record research grants received and papers published. DPI&F has had a project management system but it was no longer widely used.

RESEARCH ROLE IN TRIPLE BOTTOM LINE REPORTING

The majority of researchers felt that they had some role in Triple Bottom Line reporting as their research is linked to the impacts and the funding. This role was not in actually gathering data and assessing outcomes and impacts, it was largely viewed to be a consultative one that included providing input to what indicators to look for, suggest possible impacts and review the results of independent evaluation. Several suggested that researchers could provide outputs and contacts to people for an evaluator to meet with. Some were interested in being involved whilst others felt an obligation to contribute as the CRDC or Cotton CRC funded them.

There were three clear reasons why researchers felt that it wasn't their role to be more deeply involved (the first two were the most widely held views):

1) ***Skills*** Most researchers felt that they didn't have the necessary skills to undertake evaluation. This theme is explored further in the next section.
2) ***Time*** Researchers commented that they were already burdened with many roles other than their research and that adding another task would have to detract from the time they had to do research. A consideration in this was that evaluation had not been factored in to current research projects.
3) ***Integrity*** Independence in assessing outcomes and impacts was considered important by some.

In exploring who's role researchers felt TBL reporting should, the most common suggestion was for it to be done by an independent person or researcher with the necessary skills. A separately funded project for evaluation was required, either to contract a consultant or employ someone to carry it out. It was strongly felt that evaluation of adoption and impacts was not a suitable role for researchers. Other suggestions included that CRDC or extension staff could do it. These options are explored further in a later section.

SKILLS

A recurring theme was the need for specific skills for evaluation. For some this was a relatively new concept and they queried how it could be possible to measure things like outcomes and impact. Others had attempted some evaluation and observed that it is *"hard work when you don't know how"*. Many of these researchers had identified that a specific set of social science or economics skills were required for this work.

> *"Evaluation of outcomes and impacts requires a specific set of social science skills that are not inherently part of research teams or of researcher's own skills."*

There was discussion of the possible models for bringing in these skills - that you could either try to train researchers to be able to do 'everything' or you could leave researchers to do a small number of jobs well and bring in the required skill sets for evaluation. The attitudes of researchers towards building evaluation skills can be grouped as:

1. Bring in (and fund) specialists with the skills. Provide researchers with access to people with the skills in economics and evaluation. *Most researchers held this view.*
2. Some researchers may be interested in developing their evaluation skills if they are going to be required to report in this way. Some felt that they could contribute to gathering some evaluation data if they better understood how.
3. Some are interested in developing an understanding of what TBL is about and how it is evaluated. For some this was a keen interest and for others it was a desire to 'keep ahead of the system'.

A further discussion with one researcher highlighted value in researchers understanding what evaluation is about. That is, that researchers did not need to be trained to become evaluation experts but rather to appreciate what it is and its value. This researcher had previously been involved in DPI Victoria's evaluation training and observed how this training had led towards researchers valuing how evaluation may guide future research, rather than viewing it as a chore. This approach had a direct, personal influence on thinking about research.

DIRECTION AND GUIDELINES FOR TBL EVALUATION

It is interesting to note that several researchers expressed a need for a clear direction of how they were required to be involved in TBL evaluation. Some of this was based on a view of 'we'll do it if we need to' and some on a desire for clarification of the roles of and linkages between all players. There was a concern expressed that the current expectation for evaluation was leading towards token compliance or avoidance as it was not backed up with clarification of what was expected, how to report and where responsibilities lay.

Researchers asked for clear guidelines from the funding agencies of: How to; What information; How to generate the information and appropriate linkages for TBL evaluation. There was also call for these guidelines to include simple performance indicators that included social indicators with measures of innovation, knowledge and learning. It was suggested by one Cotton CRC program leader that this group could design an easy system for project to program reporting.

AWARENESS AND USE OF EVALUATION STUDIES AND AVAILABLE DATA

Some researchers were aware of data and reports about adoption, attitude changes and impacts of their field of research. These were mainly those working in IPM, a field of major focus for cotton research, which has had more evaluation reports and surveys undertaken and published than other fields.

Several researchers queried what data already exists and how it might be used to answer some of the questions about outcomes and impacts. They suggested that there would be value in circulating these data and reports to researchers.

A number of researchers were aware of assessments of change in pesticide usage over time and requested that this be circulated more broadly to researchers. [NB this had been published in the CRDC Annual Reports which all researchers receive (CRDC 2003a).] A small number of researchers had access to the CCA survey data about pesticide usage (a commercial report). Some also commented that they appreciated the evaluation studies that had been done by the extension network. For example, one researcher mentioned that he refers to the study of changes in industry attitudes towards IPM (Coutts, Christiansen *et al.* 2001) for general information.

There was some discussion of how researchers themselves may use evaluation data. Some researchers indicated that data about change in industry would help them to focus their future research. It was considered to be potentially helpful in thinking more carefully at the start of a project about the research in a more holistic manner. It could also help in planning possible linkages with other projects. It was suggested that mapping the impacts of a suite of related projects against a framework would help to link people and projects and to identify gaps.

Conversely, another comment was that TBL is 'just another reporting trend' that wouldn't fundamentally change the way things are done. Some commented that they have found evaluation results useful when giving presentations to diverse audiences.

PROGRAM AND PROJECT LEVEL EVALUATION

There were divided views about whether to approach evaluation at a project or program level, and to some degree this depended on the projects. Many felt that there was a need for both. Most widely it was viewed that projects could be evaluated up to the output level whilst outcomes and impacts were better identified at the program level.

It was generally felt that outcomes and impacts were best measured at the level of a program, of a suite of related projects or at the project level for particularly large projects. Another perspective was that evaluations could focus on key 'questions', drawing on relevant research projects and measures of change as needed. A program or 'project suite' basis was seen to enable a 'bigger picture' view and to identify the linkages between research projects. Particularly at the higher level it was felt that projects built on each other and that attribution would be difficult if each project aimed to evaluate impact. Measuring impact at the program level and drawing back to the contributions of each research project was seen to be both more efficient and more meaningful. Project level evaluation was seen to be able to provide more details and also to be beneficial to the researcher. Most researchers felt that it would be suitable for researchers to provide an indication of the outputs from their projects.

"We need to look at each project and how it relates to the program outcomes."

Another consideration in taking a project based approach to evaluation was the pressure that could be placed on industry with multiple surveys if each researcher needed to measure adoption and impacts. Centralising this data gathering was considered to be more efficient. It was acknowledged that a program based approach would mean more work for CRDC or someone else, but overall it was expected to be more efficient and meaningful.

"I would love to be able to have a magic figure to say how good my project is, but really we would need to look at IPM as a whole."

PROGRESSING

In addition to the key themes discussed, a number of other suggestions were made of what might help with moving forward with TBL evaluation reporting. A few suggested that a ***working group*** or steering committee for TBL evaluation be formed that included growers together with research, development and extension. This group would be able to work through TBL reporting and thinking and identify how it best relates to the cotton research programs. They could also determine key indicators to measure. On this later point it was noted that not all things need to be measured, only those that are expected to be significant.

It was suggested that a pilot or ***test case*** would be helpful to develop the process. This could select a given research area to focus on.

> *"Build on evaluation – next time it is done it will be easier as the process will be better understood."*

Others identified that a comprehensive evaluation of each research area need only be done once every few years. This timing could be done to match the reporting cycle, but 3-5 years was considered to be feasible as this was also the timeframe in which you may expect to see some change in industry. It was felt that it could be done on a rolling basis. That is, evaluate one or two major program areas each year.

Researchers were interested in being involved in planning and discussion in order to keep the work grounded and relevant and to contribute their knowledge and understanding about their work.

Whilst researchers felt they were able to contribute, they did not feel that they could complete the TBL evaluation reporting. The reasons for this are outlined in the themes about Research Role in TBL and about the Project and Program Level approaches. It is interesting to note that whilst researchers felt they were not equipped to do evaluation well, several seemed to expect that despite this they would be required to do it in the future.

There was suggestion that extension could do surveys of the regions and provide this information to the researchers and into the evaluation process. It was considered that extension officers were best situated and skilled to measure adoption and impact due to their close relations with stakeholders, their understanding of regional issues and the nature of their work.

> *"Impacts are best measured by the people who are implementing this, eg Industry Development Officers"*

It was also suggested that CRDC program staff could do much of it through reviewing project reports from a range of quarters, looking at the impacts and talking with growers in regions which have the particular problems related to that research area. It was felt that CRDC staff were in a good position to review and compare what is contained in project reports and what they are hearing from industry and from researchers.

Overall, a dominant theme was the need to bring in expertise to undertake this work. Some, particularly researchers at Narrabri, suggested a full time economist was needed again. Others suggested social science skills were needed. In either case, it was fairly unanimous that someone, other than the researchers, needed the necessary skills and dedicated time to undertake and facilitate the evaluation. Whether this was short contracted work or an employed position was not generally specified but it was felt that evaluation should be funded as a distinct project. One researcher suggested that funding a PhD study or other research project may be a suitable approach. It was suggested that if researchers developed some evaluation skills and used their observational capacity then it would not be essential for an evaluator to be involved all the way through a project.

Another suggestion was the application of a supply chain concept for viewing research. This model could investigate the contribution of knowledge generated from research into the development of industry best management practices.

CONCERNS AND ISSUES

Researchers had some reservations about evaluation and identified a few specific concerns and issues to be considered in planning an evaluation approach:

- Be careful that the evaluation is not too superficial. This was regarded as a key limitation of Benefit-Cost Analyses. One researcher developed this issue further by querying the purpose for the evaluation. He felt that if it was purely for a reporting need then a 'quick and dirty' job without too much effort would suffice. If there was genuine value that could be gained towards improving research and extension approaches then further effort would be warranted.
- Acknowledge that some projects are not well suited to measuring an impact within the time or scale of the research. For example, it was suggested that 80% of discovery research reaches a dead end.
- Recognise multiple impacts and multiple stakeholders of much research.
- There was concern about the complexity of economic analysis, in particular how to place a meaningful economic value on environmental and social issues.
- Clarify what is needed from researchers for evaluation rather than expecting it to evolve.
- Take care not to over burden the industry with too many surveys. This was particularly considered to be a threat if each researcher individually was to measure research adoption and impacts.
- Interested only if it is to be used constructively, not as a 'witch hunt'. It was considered important that the process used was about encouraging researchers so that it was not threatening.
- There was also a feeling that *'everything tends to be overestimated'*, particularly with multiple researchers all claiming the same impacts.

DISCUSSION

Limitations

A number of limitations need to be considered when interpreting these findings. Firstly, the influence of myself as the researcher. I have been working within the Cotton CRC for three years, have been based in two of the regions and so know most of the researchers well. Being an 'insider' with already established relations with the interviewees I may have received different responses and comments than an external researcher would have. Overall I felt that this was positive, the researchers were very frank and open in sharing their views. My own perspectives from having been involved with this research group and in extension and evaluation for some time will have naturally influenced the perspectives with which I explored the questions and interpreted the results. I have attempted to minimise this influence though repeatability and using systematic process in analysing the interview transcripts. For uniformity, all the interviews were conducted by myself using a standard base questionnaire. There is some inherent variability as some researchers I have worked with, or interact socially, more than others. Also, some interviews were shortened due to time constraints.

A Benefit Cost Analysis had been recently conducted, interviewing and later presenting results to many researchers. This exposure to BCA may have increased some researchers' current awareness and thinking about evaluation and the way by which they responded in these interviews. It may have also increased the focus on economic analysis and on using external expertise to conduct evaluation analyses.

Discussion of the Research

In exploring these findings a little further, Table 9 seeks to draw out if the "match" between research funders' perspectives, researchers' views, current practice in research and extension and the literature; and opportunities for progressing each of these issues.

Table 9 **Identifying the match between perspectives and practice from research managers, researchers and literature, and opportunities.**

Research Manager's Key Elements	Researchers' Perspectives and Current Practice in RD&E[1]	Literature	Opportunities
Frameworks and processes	Not used - universities have some for outputs only. Researchers have had very little exposure to these. Extension have used program logic - Bennett's Hierarchy & Logic model Unclear about what is expected of researchers. Uncertain of how evaluation could be done.	Many possible frameworks. Logic models are the simplest *p23*. Bennett's used throughout DPI Victoria, adapted for research. Adapt processes to suit *p21*. Clarify purpose and philosophy for evaluation. Consider different processes being used by other funders and agencies *p35* Evaluation strategies are best if planned in a participatory manner tailored to the program *p21*	Test possible frameworks with a few researchers and research funders. It may be practical to use one for evaluation and planning (eg Bennett's) and then integrate these findings into another for TBL reporting (eg Balanced Scorecard. Develop procedures, strategies and guidelines that clarify purpose and process. Pilot test to develop.
Need researchers to be thinking about impacts and change in industry	Some think about it in detail, some very little. Few think through the social impacts. Views about research role range from Information provider to Change agent. A high proportion of researchers are motivated by contributing to positive change in industry practice. Some have regular feedback about adoption of practices, but few have this on a whole of region or industry perspective.	A purpose for evaluation *p17* Benefit from using frameworks early. Owen's Forms 4 & 5. Philosophy of evaluation for "good thinking" *p16* Build evaluation capability and culture *p40*	Evaluation focussed on continuous improvement as well as accountability. Consider Owen's 5 forms in clarifying evaluation strategy. The values of many researchers match with thinking about impacts – the link between research and impact could be strengthened. Feedback about impact could be increased.

[1] Note some references to evaluation draw from previous unpublished work.

Research Manager's Key Elements	Researchers' Perspectives and Current Practice in RD&E	Literature	Opportunities
Processes to demonstrate outputs-outcomes-impacts	Confusion in both R&E about what each of these levels are. Researchers can readily provide outputs of research programs.	Most evaluation frameworks focus on this, particularly program logic. Cost Benefit Analyses focus on inputs, outcomes and impacts.	Clarify the interpretation of each of these terms. Develop and adopt guidelines and frameworks.
Links between research and extension	Several noted that evaluation could help to identify gaps and linkages. *Explored in greater detail in a separate study currently underway.*	This is part of using evaluation to aid project design and process. *p19* Logical frameworks try to 'force' this to some degree *p 28.* Participative evaluation planning and review has been seen to help project process and links.	Separate study will highlight opportunities to improve linkages. A participative evaluation process may help.
Gather feedback about impacts that can be attributed to research	Other than for the clear research "products", researchers were uncertain how this could be done. Some expectation for it to be quantitative. Concern about all researchers claiming the same benefits. Extension evaluations have tended not to look at attribution.	Frameworks, Program logic, Livelihoods, etc *p23.* Attribution may be difficult. Many approaches do not attempt attribution to individual projects – particularly where part of a larger, collaborative effort.	Implement systems to gather the evaluation data and a transparent, logical process to link it to research. Clarify whether attribution is needed to be to individual projects or research in general.
Independent assessment of economic return on investment	Many believed independence was important for integrity and skills. Some mention economics. Generally considered difficult unless there is a clear 'product' from the research.	Difficult to value non-market features such as environmental or social benefits – but essential *p10* - processes are being developed. *p11*	

Research Manager's Key Elements	Researchers' Perspectives and Current Practice in RD&E	Literature	Opportunities
Means to measure TBL	Need someone with the skills to do this. Researchers can contribute information, advise on indicators, review results. Some scepticism about what it means. Most readily identify economic benefits, some environmental and social – generally more difficult to understand social. Analyses presented by CRDC have aided understanding.	Cost-benefit analyses *p10.* TBL indicators *p9.* CRDC identified indicators for TBL of their research portfolio *p9* Social indicators *p12.* Frameworks *p23 ->*	Further develop the examples of TBL already done by CRDC to a few other program areas. These aid understanding and process. Identify what existing data is available for each of the key indicators developed by CRDC. Develop additional indicators if needed for CRC investments.
Groups of projects with both research and extension.	Projects need to be considered, much is easier at. Projects can contribute outputs, programs or project clusters for outcomes and impacts.	Take care to avoid double accounting *p 34* Large projects/project clusters logical for evaluation *p 35*	Identify relevant project grouping for evaluation – matched to CRDC's strategic plan outcomes would be a logical starting point.
Avoid excessive use of surveys	Survey was the main mechanism researchers considered. Concern about pressure on industry if each research independently did evaluation. Focus group and interview studies done by extension were valued by researchers. Also have comment that growers like being in some focus groups.	Literature (not included here) outlines many diverse methods for evaluation. Careful planning and processes can aid efficient use of evaluation data. Consider purpose.	Explore opportunities for non-obtrusive data gathering through regular work practice – eg recording types of phone queries, empowerment evaluation methods. Use secondary data where possible. Conduct few, well planned, integrated studies.

This table identifies a number of similarities between perspectives. There are also opportunities for moving forward and it is worth noting that there are experiences noted in the literature on which the process can draw from, and also similar challenges.

UNDERSTANDING THE CURRENT SITUATION

How do cotton researchers currently approach, think of and practice evaluation of the triple bottom line of their research?
What are their research approaches and values that may influence this?

Both the interviews and the review of existing reports confirmed that very limited evaluation currently occurs in Cotton RD&E programs, particularly of outcomes and impacts. The development of an evaluative culture and practice in cotton research is in its early stages. This is consistent with other rural research groups, many of whom are now beginning to increase their processes for evaluation of the triple bottom line impacts of RD&E investments (pg 47).

The low level of awareness, skills and understanding about evaluation processes can likely be attributed to the limited exposure that researchers have had to the concepts. A previous survey of the extension network identified a reasonable understanding of evaluation principles and frameworks but limited knowledge of and confidence in using evaluation methods. This follows a gradual building of evaluation in the extension network over two years.

Cotton researchers have a diversity of perspectives about the role of their research, their research values and the mechanisms by which they gather feedback and plan their priorities. It could be expected that these differing world views will influence how they react to evaluation. A value chain approach to research impacts may be helpful for understanding the impact of research from those who regard their role as being information providers.

The differing mechanisms used for gathering feedback and the perceptions of what constitutes success will also likely influences the way in which evaluation is perceived and utilised. For example, researchers who place highest regard on scientific elements of success have little personal need for evaluation. On the other hand, researchers who are driven by 'industry success' could be expected to be more likely to value data that demonstrates the successfulness of their research in terms of adoption and impacts. Many researchers related to a mixture of these elements of success and so any evaluation strategy would need to consider the merit of each of these criteria.

It is interesting to note that of the many researchers whose focus lay in contributing to changes in industry practice, very few had mechanisms by which to understand the level of impact their research had achieved. A reasonably high proportion readily received feedback from growers and consultants through ad hoc discussions. However, very few knew what the impact on the industry as a whole had been. Whilst this is not currently regarded as the researchers' role, it is simply curious that even of those whose personal motivation is in contributing to change, very few have the information to know how widely they have. Some appreciated of the information about impacts presented by CRDC or the extension network, and could be expected to be very responsive to evaluation that is done in a constructive manner.

Also of interest in planning evaluation approaches are the researcher's perceptions of who their stakeholders are. As so many regarded growers and consultants as key stakeholders, evaluation approaches that aim to understand and demonstrate the benefits to these stakeholders could be expected to be well worthwhile. Several also regarded the "people of Australia" as stakeholders, and this relates particularly to the environmental and social impacts. CRDC and Cotton CRC did not feature strongly as stakeholders – in part because CRDC was regarded as a representative of industry, whilst the Cotton CRC tended to be seen to represent industry and the community. The perception about the organisations as stakeholders was varied, they didn't rate as high as industry or public. Many commented that this was because their organisation's priorities fit well with industry priorities. This may need further consideration if these sets of priorities were to grow apart in the future.

Researchers are passionate about their work. They do it because they enjoy the science and quest for knowledge and/or because they have a commitment to contributing towards improvements in industry practice. Several expressed concerns about balancing organisational and industry expectations. Any evaluation strategies will benefit from engendering ownership amongst industry and research organisations as well as researchers.

Evaluation processes will do well by clearly contributing towards, rather than competing with, researchers core values and interests. It needs to be seen as something that can contribute to researchers' work, and in turn lead to peer recognition, research funding and progression. Without this, there is risk that evaluation may be perceived as more bureaucracy that simply interferes with the job of doing science. This could in turn lead to token compliance rather than learning and improvement.

MOVING FORWARD

What strategies may be used to move the triple bottom line evaluation approach forward?

This research has presented the diversity of expectations, understanding and skills of the researchers, and the processes that they suggest would help for progressing TBL evaluation. Research managers indicated that there is a strong need for accountability. They also wished for systems that would contribute to improving research by strengthening links between research and extension and by encouraging researcher to think about impacts of their research. Researchers themselves also stressed that evaluation needed to be approaches in a constructive rather than fault finding manner.

CRDC's existing practice of presenting to researchers some worked examples of the types of indicators, how they may be measured and some results has been helpful in building researchers awareness and appreciation for TBL. Several mentioned the usefulness of the various graphs on industry changes attributed to IPM practice that CRDC had compiled (CRDC 2003a).

Many researchers indicated a willingness to be involved in evaluation of research programs. This interest was strongest amongst those whose values and views about research were linked towards facilitating change in the industry. Researchers, research managers and the literature highlighted the importance of an evaluation ethos that was constructive rather than fault-finding. Developing and clearly communicating this ethos will enable more productive approaches to TBL evaluation in cotton RD&E.

As discussed on page 29, Wissemann (1992) identified that a first step in planning evaluation is to clarify the purpose for it. The interviews also highlighted a strong need for clarity about evaluation in Cotton RD&E, and what is expected from various stakeholders. The call for direction, guidelines and skilled expertise to guide evaluation has identified an important first step.

A small working group, as suggested by some researchers, would be suitable to progress evaluation and develop these guidelines. This working group would ideally involve representatives from research, extension, CRDC, Cotton CRC and researchers agency managers. The groups will be able to take into account the research and literature perspectives presented in this study, and the evaluation skills and understanding of the extension network (Christiansen 2003, unpublished survey).

There is a clearly identified need for a strategic, holistic view to evaluation in order for it to be efficient and to minimise the pressure on industry in gathering data. This efficiency has three core elements. The first element is in the balance between projects and programs. The general view is that projects do need to be looked at individually, up to the level of outputs. For higher level outcomes and impacts, the evaluation would be best to focus on research programs, clusters of projects or key evaluation questions. For example, the "outcomes" identified in CRDC's Five Year Strategic Plan (CRDC 2003b) could be used as the key questions for TBL evaluation.

A second element of evaluation efficiency is in careful planning and use data collection methods to minimise the pressure on industry through evaluation. Sharing evaluation studies across programs and utilising non-obtrusive means to gather data can aid this. For example, secondary data can be utilised (eg CCA's market report for chemical usage; Workplace Health and Safety Statistics, River Health Monitoring, etc) and data may be gathered through existing research and extension practices (eg empowerment evaluation in training courses; researchers recording the nature of the queries they receive). Some smart thinking in IT systems may also aid in gathering evaluation data non-obtrusively.

A third element is to develop an efficient evaluation strategy. Such a strategy would to clearly identify the roles and responsibilities of all players, the key skills, a lead contact for data gathering and reporting, purpose and use and the resources available. It would also recognise that not all projects need to be evaluated (as outlined on pg 22), and identify the criteria for choosing when to evaluate.

It is interesting to note that many researchers considered that evaluation requires a specialised set of skills. Whilst many were unclear of how adoption or impacts may be measured, if at all, there were also many who could see that it would be possible but didn't feel that they had the necessary skills.

Almost unanimously, researchers did not want to be the people actually doing the evaluations. They were willing top be involved but due to time, skills, and / or integrity felt that they were not well suited to gathering and analysing the evaluation data. Many researchers suggested that evaluation needed to be managed and resourced as a special project that brought in the required expertise and the dedicated time to work on it. This matches with the recommendations of (AACM-International 1997). As many researchers were interested in being involved, and the researcher managers wanted researchers to develop their understanding of impacts, a suitable model may be one that mixes internal and external capability, as described by (McDonald, Rogers *et al.* 2003).

A concern about upskilling researchers to "do all jobs" rather than to focus on their science appeared to reflect broader concerns. There was a general feeling that the researchers' capacity to do science was being diminished due to the time that required to do various other tasks in research management, extension, reporting and bureaucracy. This concern will need to be carefully considered by the CRDC and Cotton CRC in planning an evaluation approach.

As described by (McDonald, Rogers *et al.* 2003), an optimal model for investment in evaluation needs to balance the internal evaluation capability with external reporting needs. Researchers identified that it needs to consider: the time demand on researchers; the relative effectiveness of using internal or external capacity; independence; integrity; and skills.

For evaluation strategies to be valuable, it will be important to engage stakeholders in the development of an evaluation process for Cotton RD&E. By conducting interviews with many of the researchers, this study has commenced that engagement process with the research sector. The extension network have previously been closely involved in the development of evaluation approaches (Christiansen, Pyke *et al.* 2003). Further engagement of the research users, funders and agencies would be helpful. Other mechanisms for participation suggested by researchers included an evaluation working group, forums to discuss findings and to develop and test evaluation processes by piloting it with a few research programs or project clusters.

Several researchers called for evaluation processes to be interactive. Whilst not regarding it as their role or skills to conduct the evaluation, many researchers were keen to have involvement in guiding and informing the evaluation. There was also suggestion that bringing together key researchers, perhaps with industry, to discuss initial evaluation findings would be a beneficial process. Involving both research and extension in this process may help to build the linkages between these two sectors, which many researchers consider to be weak (Christiansen 2004, unpublished). This may also help to increase frank discussion and constructive feedback about science, that some considered were lacking.

The views of cotton researchers are reinforced by experiences in Victoria (McDonald, Rogers *et al.* 2003; pg 40) and previously in DPI&F that specialised expertise and capacity are needed to drive the evaluation process. Research managers have identified that they wish to build the habit of researchers in thinking about outcomes and impacts. They have also identified a need for some external assessment of impacts. In this way, all elements of this study suggest that a suitable model would be one that involves researchers in evaluation and builds their appreciation of its value whilst bringing in a third party to design, conduct and report TBL outcomes and impacts. Suggestions are that this third party may be in-house or external, but that they do need to interact with researchers at several stages of planning, data gathering and interpretation.

Both literature and interviews indicate merit in building researchers' appreciation for the value of evaluation. This appreciation is particularly important to overcome the concerns that are discussed on page 38 (Patrick and Niles 1988) and raised in the interviews (pg 90). Upskilling researchers to understand how evaluation is done may also lead to researchers gathering useful data during their research. This may help provide useful data for evaluation, which has often posed a difficulty as described on pg 37 by (Gordon 1998) and (Kingsbury 2003).

Multiple approaches considered important to gradually build evaluation capability (pg 40; pg 80). McAllister (1998) suggests that building these skills, processes and thinking can take 8-10 years. However, there is need for some clear decisions and communication between research funders and researchers about what is expect and useful in terms of evaluation.

Several researchers indicated that a pilot study with a few programs or projects would be helpful. These pilots would help to develop an evaluation strategy which clearly outlines the processes and expectations for evaluation in Cotton RD&E. Ideas for this are progressed in more detail in the next section.

There is interest in utilising evaluation to improve research and extension practice. A large body of literature exists about learning organisations and continuous improvement processes. These concepts relate to the ideas about the use of evaluation and feedback for improving or focussing research. It would be worth further exploring these opportunities, and understanding how the role of the culture of the various organisations (research funders and research agencies) may impact on how evaluation is adopted and utilised.

The researcher managers expressed a need for frameworks for RD&E evaluation. Currently, such frameworks are not used in any of the research groups interviewed. Researchers had very limited awareness of such frameworks. One researcher suggested that frameworks would only be of value for researchers who are educated in understanding evaluation. The frameworks may provide a useful basis for building the understanding about how to evaluate and report on TBL research outcomes and impacts. This has been the case in developing an understanding of evaluation in the extension network. DPI Victoria initially avoided introducing set evaluation frameworks to researchers (McDonald, Rogers *et al.* 2003). However, they found that, left to their own devices, researchers took a long time and developed far too complex models.

Using an existing framework helps to reduce some of this frustration, reduce time taken, ensures that the framework matches reporting requirements. There are a suite of tried and proven frameworks (described pp 23-33) that Cotton RD&E could select from and modify as needed. It may prove functional to utilise one framework for project planning and evaluation (eg Bennett's Hierarchy) and then for the research funders to integrate these measures into another framework for TBL reporting (eg Balanced Scorecard or Logic model). Various frameworks could be tried during the pilot phase. Alternatively, the extension network has already begun to use the LOGIC model and Bennett's Hierarchy. These simple frameworks can be readily utilised across a range of project types.

Effectively using evaluation can be a problem, as identified in the literature (pg 36 and pg 50) and in the concerns of researchers. If researchers perceive that evaluation is initiated but then not used, they will likely become negative towards it, as some already are towards the annual reports.

In addition to demonstrating the impact of their research investments, research programs also face a greater challenge of enhancing this impact. Whilst TBL evaluation is currently driven by a reporting need, there is an opportunity for evaluation processes to be developed in a way that increases the impact of research into practice. As described by Horton and Mackay (2003; pg 7), evaluation can be an opportunity to aid agricultural research communities in responding to change.

Research managers also identified that evaluation processes could be used for both accountability and continuous improvement. In a similar vein, most researchers' prefer systems that are of greater benefit than solely a reporting need. There was a feeling that discussion in the evaluation process could help to strengthen linkages, review progress, identify gaps and develop future research needs.

As identified by one researcher, this purpose for evaluation will dictate the level of investment that goes into it. He suggested that if purely for reporting, a 'quick and dirty' job would suffice, but that if greater benefit can be gained then it would be worthwhile to put in more effort. Evaluation systems will need to be adequately resourced to be able to achieve these multiple goals. This investment includes researchers' time, users time, travel, meetings and evaluation expertise. Researchers suggested a specific evaluation project that commenced with a pilot program would be suitable.

Clarifying evaluation use will also determine the level of investment and focus on economic assessment of the TBL of research investments. The difficulty of placing economic values on social and environmental impacts was widely acknowledged both in the literature (pg 10) and through the research findings. There is a risk that the entire evaluation process may be poorly perceived if it is focussed on economic valuation, which researchers tend to be highly sceptical about. This risk will need to be managed through careful balance and transparent processes. Clarification of the need for economic assessments or the potential to use other indicators will aid in determining future strategies.

OTHER ISSUES

In addition to the key research questions about evaluation, a number of other issues and opportunities have been highlighted, as described below.

Many researchers would appreciate more detailed feedback about their research from the funding agencies and for some, interactively from other scientists. This related particularly, but not exclusively, to the final reports. Short loop feedback between scientists and industry occurred readily for the Narrabri and Toowoomba based researchers but less so for university based researchers. However, some researchers felt they had little in-depth feedback from CRDC or other researchers. There was added concern that feedback and collaboration across organisations may become restricted in the future as funding shortages made research more competitive.

Researchers' perceptions about the role of their research and their core values frequently revolved around change in industry. Extensive literature surrounding participatory research approaches focus on this role of research. Whilst many researchers held values and approaches that aligned with those of participatory research, this school of though is rarely referred to in cotton research programs. It may be worth exploring whether participatory and action research theories may help some cotton researchers to develop valid research approaches that assist this ideology of science into practice.

There is a need to clarify the expectations that funding agencies have of researchers. One element already discussed is evaluation, another is scientific publishing. Several researchers felt that publishing scientific papers was not important to the industry. They felt that both ACGRA and CRDC were interested only in communication with industry and were not interested in whether scientific papers were published. This was seen to conflict with organisational expectations for publishing. In further exploring this issue, CRDC commented that they considered it important for the science that they fund to be published in journals.

> *" We'd also like some researchers to publish a bit more frequently."*

This needs to be clarified with researchers, and embedded in any evaluation approach. Some mentoring or training for researchers of how they can balance these various elements would be valuable.

Linked to the above issue was that of incentives for researchers. Perceived differences between organisational requirements for salary rewards and the funding agency's expectations were a source of discontent and stress for some researchers. Clearer communication lines may help to resolve this, though there may also be a need to review the incentive structures.

The concern about the lack of training provided for researchers in the diverse aspects of their roles may warrant further investigation and consideration.

> *"I am appalled by the lack of training I receive in things such as project management and project evaluation."*

Whilst many researchers hold PhD training in the science discipline in which they specialised and some short training in areas such as media skills, many felt that they also require other skills. They felt that they had limited opportunities or encouragement to develop skills in project management, extension or evaluation.

CONCLUSIONS

There is currently limited knowledge about or practice of evaluation amongst cotton researchers. However, as many of the researchers values focus on change in industry, there is a good base to start from. Cotton researchers are a diverse group who are passionate about their science, and not wanting to be challenged with additional roles for which they don't feel skilled.

There is a need for clarity in the interpretations of the levels of outputs, outcomes, impacts. Both research and extension groups have diverse interpretations of these terms, and multiple approaches are used. Clarification is also needed of the expectations of researchers in the evaluation process.

There is a clear need to build evaluation capability if evaluation is to occur in Cotton RD&E. Both cotton researchers and other experiences described in the literature prefer a model of bringing in evaluation expertise to drive the evaluation process, whilst the researchers are involved with the process and build a general awareness and appreciation for its value.

Development and pilot testing of an evaluation strategy though a participative process is recommended.

FUTURE RESEARCH AND DEVELOPMENT

This research has identified a number areas for further exploration, including issues relating to evaluation and more broadly in research process and management. The current study has focussed on the perspectives of one stakeholder group, the researchers. Further studies could explore the perspectives of other stakeholders such as end-users, research managers, research agencies and politicians. Whilst previous studies have worked with the extension group, they have not researched the same issues in depth so some further perspectives from extension could be gained. Some perspectives of research funders are included here. However, it would be helpful to follow on from this study to gain more clarity on the essential needs from those funding agencies in time frames, resources, staff and perspective on purpose. In planning the overall evaluation approach, it could be beneficial to seek advice from DPI Victoria's evaluation unit and others who have had previous experiences in developing evaluation capability. There may be potential for this unit to deliver evaluation training for Cotton RD&E staff.

A primary area for future research in evaluation in cotton research programs would be to develop and conduct a pilot case with a small number of research programs. A number of elements will need to be further developed to contribute towards these pilot studies, as illustrated in Figure 10. The development and learnings from these pilots could then be used to build an evaluation capability strategy for Cotton RD&E.

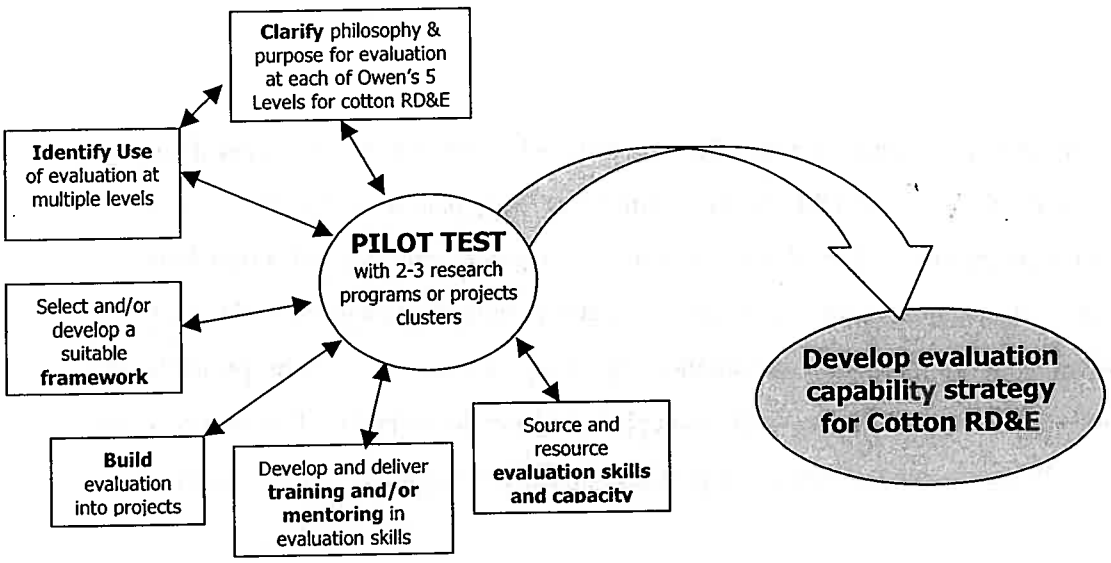


Figure 10 **Potential elements of a pilot research and development program for building evaluation in cotton RD&E programs.**

In implementing these pilot studies it would be valuable to research the impact of this process on participants' attitudes and approaches towards evaluation, extension and their research process. This could involve a longitudinal study that explores the influences of evaluation processes on researchers values and approaches over time. It would be valuable to involve both research and extension in this study, and to explore how the pilot process contributes towards linkages and mutual understanding between research and extension.

A detailed study could go further to explore the potential for evaluation to lead help in building learning organisations. Drawing on the organisational behaviour and change management theories, this research would need to look in greater depth at the individual organisations and their respective cultures and approaches to innovation. It could explore and develop the dynamic between organisations, funding agencies and cooperative research centres.

This pilot could be conducted for 2-3 programs or project clusters. Different approaches could be used in each in order to compare and contrast the participants' reactions to the program with or without certain elements. Alternatively, pilots could be developmental with each building on the experiences and understandings gained from the previous pilot.

The time period and detail of these pilots can be developed to suit the necessary timeframe and budget. For example, DPI Victoria's initial evaluation pilot (of Target 10 – a dairy extension project) was able to draw on previous cost-benefit analyses and comprehensive databases of activities (McDonald, Rogers *et al.* 2003). Within two months of the pilot, project staff produced their first evaluation report. By the end of 18 months, principles and a framework for developing an evaluation plan had been developed for this project. Over the next 18 months, an evaluation support team tried out evaluation with volunteer projects.

Based on these experiences, and the work already done through this project, a concentrated effort to pilot evaluation with projects clusters or programs in cotton could be done initially over a few months. By building on these learnings, together with the findings of this study, an evaluation plan could feasibly be developed within six months. This would rely on cooperation of research groups, access to data and dedicated time from one or a few people to drive the process and gather the data.

Through these pilots it would also be important to work with eth research agencies to ensure synergy between the evaluation approaches of Cotton RD&E and the merging needs of the agencies. For example, under DPI&F's new structure, evaluation will be required from project, program and outcome levels.

Other issues for further research that can be drawn from this study include:

1. Explore further the issues and opportunities relating to the need for support, mentoring and training in the range of skills that researchers require. These include areas such as project management, leadership, extension, time management, writing and in balancing some apparently conflicting demands.
2. Identify opportunities for researchers to contribute evaluation data through simple observation and recording of the issues and trends that they observe in their regular interactions with industry.
3. Explore opportunities for research incentives to reflect the growing importance of communication with industry and demonstrated impact.

4. Identify the potential for participatory research and development approaches to enhance the understanding and application of differing roles for science.

REFERENCES

Literature

AACM-International (1997) 'Guidelines for the selection of Evaluation Techniques to Assess R&D Programs.' Land and Water Resources Research and Development Corporation, IR01/97, Canberra.

Adato M, Meinzen-Dick RS (2002) 'Assessing The Impact Of Agricultural Research On Poverty Using The Sustainable Livelihoods Framework.' International Food Policy Research Institute, EPTD Discussion Paper 89.

ANAO (2002) 'The Future Direction of Audit – A National Audit Office Perspective.' Australian National Audit Office, Commonwealth of Australia.

Armstrong A, Francis R (2003) Social indicators - promises and problems: a critical review. *Evaluation Journal of Australasia* 3, 17-26.

Australian_State_of_the_Environment_Committee (2001) 'Australia: State of the Environment 2001.' (Canberra: CSIRO Publishing)

Baker E (2004) A critique of the Triple Bottom Line approach. In 'Australasian Evaluation Society Lunchtime Seminar'. Brisbane

Batterham R (2000) 'The Chance to Change.' Discussion Paper by the Chief Scientist, Australia.

Bauer M, Kirchner A, Humphries J, Bueren Mv, Gordon J (2003) 'Evaluation of the Agroforestry and Farm Forestry Program. An assessment of benefits - Stage 2.' Centre for International Economics *for* Rural Industries Research and Development Corporation, Canberra.

Bennett CF (1977) 'Analysing impacts of extension programs.' (Washington DC Extension Service, US Dept of Agriculture.)

Boleman C, Cummings S (2002) Tips on Evaluation Types and Methods - An Introduction. In. (Texas Cooperative Extension, The Texas A & M University System.)

Brown R (1979) Key issues in evaluating student affairs programs. In 'Evaluation in Student Affairs'. (Ed. G Kuh) pp. 13-31. (American College Personnel Association: Cincinatti)

BRS (2001) 'Country Matters: Social Atlas of Rural and Regional Australia.' Bureau of Rural Sciences, Department of Agriculture Fishieries and Forestry, Canberra.

Carney D (2002) 'Sustainable Livelihood Apporaches: Progress and Possibilities for Change.' (Department for International Development: Toronto)

Checkland P (1984) 'Systems science and world order.' (Oxford, Pergamon)

Christiansen I (2003) 'Cotton Extension Outcomes - Collated evaluation reports.' (Cotton CRC)

Christiansen I, Pyke B, Gibb D, McIntyre G (2003) Ever improving: Evaluation and outcomes in the National Cotton Extension Network. In 'APEN 2003 National Forum'. Hobart. (Ed. F Vanclay). (The Regional Institute)

CIE (2000) 'Evaluation of the Prospective New Industry Program Project evaluations — stage 2.' Centre for Intenational Economics *for* Rural Industries Research and Development Corporation, RIRDC Publication No 00/115, Canberra.

Coutts J, Christiansen I, CottonExtensionNetwork (2001) 'Changes in Attitudes to Integrated Pest Management in the Cotton Industry: 1997-2001 and Attitudes to Area Wide Management. A Focus Group Study.' CRDC, REC-UQld and Cotton CRC, Narrabri.

CPA_Australia (2004) About triple bottom line reporting: What is triple bottom line? In. (Ed. COI Centre).

CRDC (2003a) 'Annual report 2002-03.' CRDC, Narrabri.

CRDC (2003b) 'Five year strategic research and development plan 2003-2008.' Cotton Research and Development Corporation, Narrabri.

Dart J, Petheram RJ, Straw W (1998) 'Review of evaluation in Agricultural Extension.' (RIRDC Publication No. 98/136)

Davies I (1999) Evaluation and Performance Management in Government. *Evaluation* **5**, 150-9.

Dick B (1993) 'You want to do an action research thesis? How to conduct and report action research.' (Department of Psychology, The University of Queensland.: St Lucia, Brisbane, Australia)

Dick B (1997) The Snyder evaluation process: an overview. In 'Resource Papers in Action Research'. (Southern Cross University)

DNR&E (2001) 'Innovation: Sustaining Export Growth in the Food and Agriculture Sector.' Department of Natural Resources and Environment, Government of Victoria, Melbourne, Victoria.

Dumsday R, Sturgess N (2003) 'Non-Market Valuation and Holistic Assessment Part 1: Non-Market Benefits of Research and Development.' URS Sustainable Development for Land and Water Australia, Richmond, Victoria.

Ekboir J (2003) Why impact analysis should not be used for research evaluation and what the alternatives are. *Agricultural Systems* **78**, 166-184.

Elkington J (1997) 'Cannibals with forks.' (New Society Publisher: British Columbia, Canada)

English B, Kaleveld L (2003) The politics of program logic. *Evaluation Journal of Australasia* **3**, 35-42.

Environment_Australia (2003) Triple Bottom Line Reporting in Australia - A Guide to Reporting Against Environmental Indicators. In 'Sustainable Industry, Financial Services Sector'. (Australian Government, Department of Environment and Heritage: Canberra)

Fetterman DM (1996) Empowerment evaluation: An Introduction to Theory and Practice. In 'Empowerment Evaluation: Knowledge and Tools for Self Assessment and Accountability'. (Eds DM Fetterman, SJ Kaftarian and A Wandersman) pp. 411. (Sage Publications: California, USA)

GetInvolved (2004) Community Engagement Evaluation Strategy. In. (Community Engagement Division, Department of Communities, Queensland Government: Brisbane)

Gordon J (1998) 'Benefit-cost analysis of RIRDC's emerging industries program.' Centre for International Economics *for* Rural Industries Research and Development Corporation, RIRDC Publication No 98/104.

Gordon J, Davis L (1999) 'Impact Evaluation of RIRDC's Established Industries Program. An Overview of All Projects - Stage 1.' Centre for International Economics *for* Rural Industries Research and Development Corporation, RIRDC Publication No 99/90, Canberra.

Grasso PG (2003) What Makes an Evaluation Useful? Reflections from Experience in Large Organizations. *The American Journal of Evaluation* **24**, 507-514.

GRDC (2002) 'Annual Report 2001-02.' Grain Research and Development Corporation, Canberra.

Grob GF (2003) A Truly Useful Bat is One Found in the Hands of a Slugger*1. *The American Journal of Evaluation* **24**, 499-505.

Hassall&Associates (2004) 'Evaluation of GRDC Supported Farming Systems Projects.' Hassall and Associates Pty Ltd, for Grains Research and Development Corporation, Dubbo.

Hayward S (2002) The New Corporate Balance Sheet: Black, Red - and Green. *AEI Environmental Outlook.*

Henderson T (2003) A systems approach to the evaluation of collaborative R&D projects in the Australian sugar industry. In '1st Australian Farming Systems Conference'. Toowoomba. (Eds B Robinson and C King). (The Australian Farming Systems Association Inc.)

Henry GT (2000) Why not use? In 'The expanding scope of evaluation use. New directions for evaluation.' (Eds V Caracelli and H Preskill) pp. 85-98. (Jossey-Bass: San Fransisco)

Horton D (1998) Disciplinary roots and branches of evaluation: some lessons from agricultural research. *Knowledge and Policy* **10**, 31-66.

Horton D, Mackay R (2003) Using evaluation to enhance institutional learning and change: recent experiences with agricultural research and development. *Agricultural Systems* **78**, 127-142.

Kaplan R, Norton D (1996) 'The Balanced Scorecard: translating strategy into business action.' (Harvard Business School: Boston)

Kaplan RS, Norton DP (1992) The Balanced Scorecard - Measures that Drive Performance. *Harvard Business Review.*

Kellogg-Foundation WK, Sanders J (1998) 'W.K. Kellogg Foundation Evaluation Handbook.' (Collateral Management Company: Battle Creek, Michigan, USA)

Kingsbury N (2003) 'An evaluation culture and collabroative partnerships help build agency capacity.' Office of Management and Budget, USA, Washington DC.

Kolb DA (1984) 'Experiential Learning: Experience as the Source of Learning and Development.' (Prentice Hall: Englewood Cliffs, New Jersey)

Lawrence DN (2002) 'Developing besst research/extension practice through evaluation of process and impact.' QDPI.

Lawrence DN, Carberry PS, Cawley SC, Doughton JA, Herridge DF (2001) Success factors for participatory farming systems projects. In '10th Australian Agronomy Conference'. Hobart

Lawrence DN, Christodoulou N, Kuskie J (2003) A Framework to Plan and Evaluate Activities in Participatory Grains Research, Development and Extension (RDE). In '1st Australian Farming Systems Conference'. Toowoomba. (Eds B Robinson and C King). (The Australian Farming Systems Association Inc.)

Leeuw F, Toulemonde J, Brouwers A (1999) Evaluation activities in Eurpoe: A Quick Scan of the Market in 1998. *Evaluation* **5**, 487-496.

Lubulwa G (1996) An economic evaluation of the impact of ACIAR funded agricultural research based in Africa. In 'Annual Conference of the Australian Agricultural and Resource Economics Society'. Melbourne

Mackay R, Horton D (2003) Expanding the use of impact assessment and evaluation in agricultural research and development. *Agricultural Systems* **78**, 143-165.

Mark MM, Henry GT (2004) The mechanisms and outcomes of evaluation influence. *Evaluation* **10**.

McAllister E (1998) Overview and conclusions. In 'Selected Proceedings of the World Bank Seminar: Public Sector Performance - The Critical Role of Evaluation'. (Ed. K Mackay) pp. 121-5. (World Bank: Washington DC)

McDonald B, Rogers P, Kefford B (2003) Teaching People to Fish? Building the Evaluation Capabilty of Public Sector Organisations. *Evaluation* **9**, 9-29.

McDonough W, Braungart M (2002) Design for the Triple Top Line: New Tools for Sustainable Commerce. *Corporate Environmental Strategy* **9**, 251-258.

McGee P, Charlton M, Cattanach G (2002) 'Innovating Rural Australia: Research and Development Corporation Outcomes.' (Agriculture Fisheries and Foresty Australia: Canberra)

Monitoring-and-Evaluation-Working-Group (2004) 'Resource Condition Indicators Overview.' Department of Agriculture Fisheries and Forestry, Canberra.

Murray P (2000) Evaluating participatory extension programs: challenges and problems. *Australian Journal of Experimental Agriculture* **40**, 519-526.

National-Resource-Centre-for-Consumer-Participation-in-Health 'An evaluation of the National Resource Centre for Consumer Participation in Health.' www.participateinhealth.org.au.

NovoNordisk (2002) Reporting on the Triple Bottom Line. In.

NRM-Ministerial-Council (2003) 'National Natural Resource Management Monitoring and Evaluation Framework.' Department of Agriculture, Fisheries and Forestry.

Owen J, Lambert F (1998) Evaluation and the information needs of organisational leaders. *American Journal of Evaluation* **19**, 355-365.

Owen J, Rogers P (1999) 'Program Evaluation: Forms and Approaches.' (Allen and Unwin Pty Ltd: Sydney)

Owen JM (1993) 'Program evalution.' (Allen and Unwin: St Leonards, NSW, Australia)

Owen JM (2003) Evaluation culture: a definition and analysis of its development within organisations. *Evaluation Journal of Australasia* **3**, 43-47.

Patrick J, Niles S (1988) Estalishing accountability and evaluation procedures in student affairs offices. *NASPA Journal* **21**, 291-296.

Patton MQ (1978) 'Utilisation focussed evaluation.' (Sage Publications: Beverly Hills, CA)

Patton MQ (1997) 'Utilisation focussed evaluation.' (Sage Publications: Beverly Hills, CA)

PlanningNSW (2003) 'Community Engagement Handbook.' (PlanningNSW: Sydney, Australia)

Qld-Government (in press) 'Engaging Queenslanders: evaluating community engagement - DRAFT.' Community Engagement Division, Queensland Department of the Premier and Cabinet.

REC (2003) 'Research methodolgies in management and extension. Learning Guide.' (Rural Extension Centre, The University of Queensland: Gatton)

Rockwell K, Bennett C (1995) Targeting Outcomes of Programs (TOP). In. (University of Nebraska - Lincoln; Cooperative State Research Education and Extension Service, USDA)

Roth GW (2003) Measuring the sustainability of cotton production systems. In 'World Cotton Research Conference - 3'. Capetown, South Africa. (In Press)

Salvaris M (2000) 'Community and Social Indicators: How Citizens Can Measure Progress - An overview of social and community indicator projects in Australia and internationally.' (Institute for Social Research, Swinburne University of Technology: Hawthorn, Victoria)

Scriven M (1991) 'Evaluation Thesaurus.' (Sage Publications: USA)

Scriven M (1996) Evaluation: the state of the art. *Evaluation New and Comment* **5**, 5-9.

Sloane P (1999) '*Ex-Ante* Evaluation of Selected Research Projects 1997.' Sloan, Cook and King Pty Ltd *for* Land and Water Resource Research and Development Corporation, IR01/98, Canberra.

Stewart R, Mohamed S (2001) Utilising the balanced scorecard for IT/IS performance evaluation in construction. *Construction Innovation* **1**.

Taylor-Powell E (2001) The Logic Model: A Program Performance Framework. In 'Third annual University of Wisconsin-Extension, Cooperative Extension program evaluation conference June 18-21, 2001'. Madison, Wisconsin. (University of Wisconsin - Extension)

Templeton D (2004) Outcomes: Evaluating agricultutral research projects to achieve and to measure impact. In '3rd Crawford Fund/UNE Master Class in Research Management in Agriculture'. Sydney. (Ed. I Metcalfe) pp. 121 - 133. (UNE)

UN (1992) 'The Rio Declaration on Environment and Development.' United Nations Conference on the Environment and Development, Rio de Janerio, Brazil.

UNEP (1999) 'United Nations Environment Program Annual Evaluation Report.'

UNEP (2002) 'Summary report of the high-level roundtable on cultural diversity and biodiversity for sustainable development.' Governing Council of the United Nations Environment Programme.

Williams A, Williams J (2000) 'Australian Cotton Industry Best Management Practices Manual.' (Cotton Research and Development Corporation: Narrabri)

Wissemann A (1992) 'Three Key Concepts in Extension Evaluation.' Research and Extension Unit, Qld Department of Primary Industries, Brisbane.

World-Commission-on-Environment-and-Development (1987) 'Our common future.' (Oxford University Press: Oxford, UK)

Personal Communications

Jessica Dart, Evaluation and Training Consultant.

John Dixon, Food and Agriculture Organisation, Rome.

Peter Ellsworth, IPM Specialist, University of Arizona.

Rob Fearon, CEO, Coastal Zone Cooperative Research Centre.

Gus Hamilton, Chief Extension Officer, QDPI&F

Tracey Henderson, Program Manager, Sugar Research and Development Corporation.

David Lawrence, Project leader, Eastern Farming Systems Project, Qld DPI&F

Bruce Pyke, Research Manager, Cotton Research and Development Corporation

Arnold Wissemann, Principal Policy Officer, Office of the Chief Scientist, QDPI&F

Appendix 1 Research Project Proposal

Evaluation for Continuous Improvement: the application of a conceptual framework to evaluate and identify gaps and opportunities at a program level.

Case Study: Cotton Research Programs – 2 areas (eg IPM; Soils)

Research Approach: Constructivism

Bodies of Theory: Evaluation; Research Management; Business

Key Questions: Is there potential to improve research and extension process and linkages in the cotton industry through using evaluation frameworks?
- If so, can the process meet the need for Triple Bottom Line Reporting?

Background

"Over the years the United Nations Environment Programme (UNEP) has learnt... that evaluation is not a fault-finding exercise but rather a strategic management tool that allows management and programme staff not only to review progress made but also to critically reflect on the constraints and challenges of delivering a quality global environmental programme".

Dr. Klaus Topfer, UNEP Executive Director, 1999 Annual Evaluation Report.

Research agencies are under increasing pressure to demonstrate triple bottom line impacts of research and extension investments. Fundamental to achieving this is the need for a process to evaluate outputs, outcomes and impacts. Evaluation frameworks applied at a research program level have a potential to both measure impacts of research and to identify opportunities for enhanced linkages or new project priorities. Evaluation at a program (rather than project) level is also generally less threatening for individuals, builds a greater picture of overall impact and is not focussed on attribution to individual projects.

Within the Cotton CRC extension network, evaluation has been approached with the dual purposes of 1) review and measuring impacts and 2) for planning and continuous improvement. Over the past 3 years we have gradually built the evaluation skills and outcome focus of the cotton extension network. There has been relatively focus on evaluation of research programs though this is changing as CRDC introduce more evaluation requirements to their project reporting formats.

Research Plan

This project will investigate whether a conceptual framework for evaluation across research and extension programs can help to enhance linkages and outcomes. It is thought that a focus on measurable impacts could help research and extension to develop shared outcomes and to clearly identify the contributions of various players in the research-extension-practice loop. It will also investigate whether these frameworks can be of help in reporting on the impacts for industry, environment and community of research and extension investments.

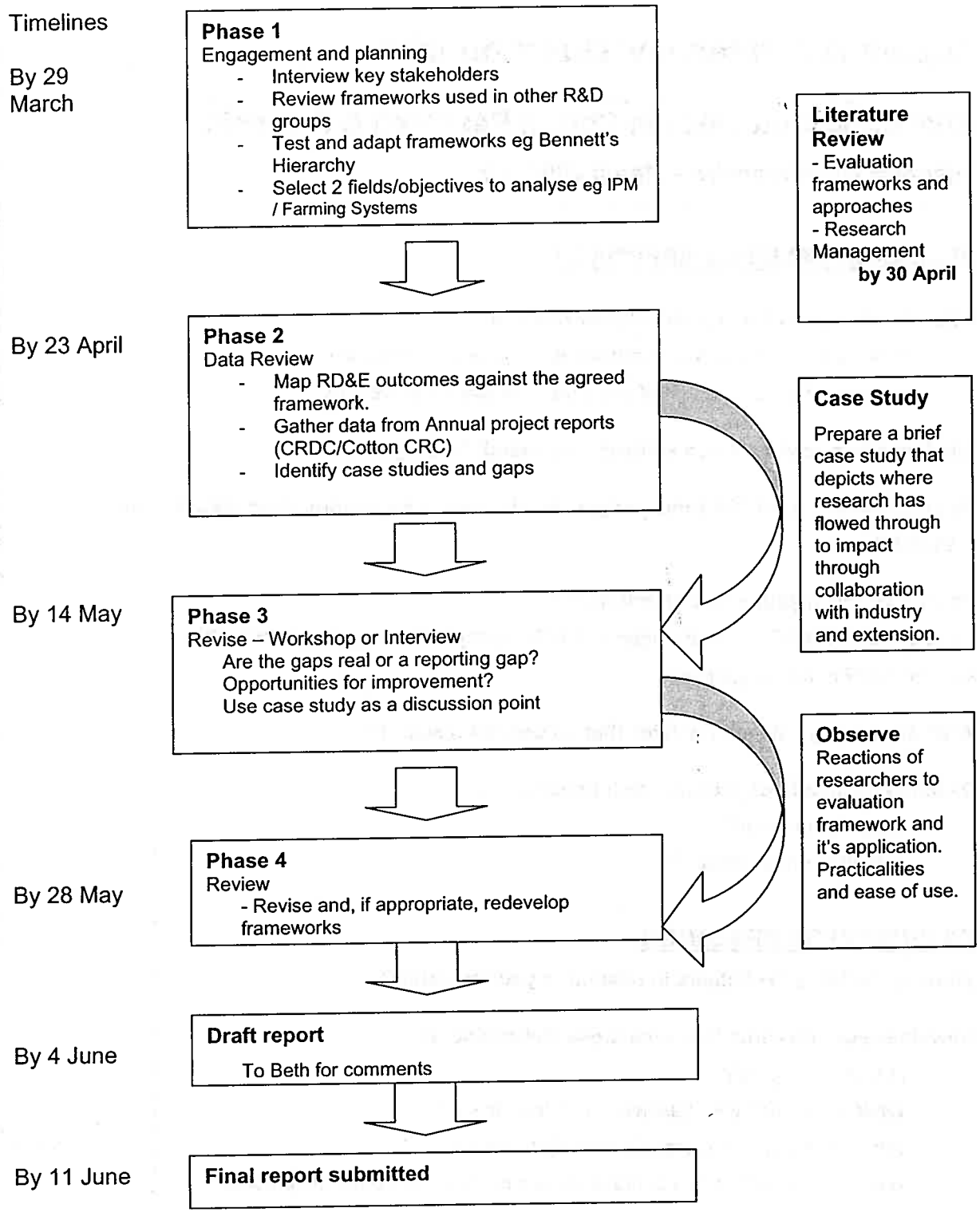
Bennett's Hierarchy and the Logic model have been used within the cotton extension network due to their great flexibility for a range of projects and simplicity. A modification of these frameworks will be tested for usability across cotton research programs through the research detailed in Figure 1. The detailed research approach will be developed in Phase 1 through consultation with key stakeholders. Data will be gathered initially from projects reports and later reviewed through interviews or workshops with researchers. It will be reviewed from an overall evaluation perspective as well as whether it can identify triple bottom line outcomes.

The research will include observations of the reactions of researchers and funders to the evaluation framework and identify whether the data has helped to identify areas for improvement in research-extension collaboration, planning and management.

Resources

Reports of various evaluations of extension programs.
CRDC and Cotton CRC Annual research/ project reports.
Annual reports
Direct contact with research and extension staff.
Travel and workshop budget from existing CRC/CRDC project.

Figure 1: Research Plan



Appendix 2 Interview Questionnaire

Measuring Outcomes in Cotton Research & Extension

Interview Questionnaire – March 2004 v 2

RESEARCH VALUES & APPROACH

What do you see as the key role of your research?

- *How does your research fit with influencing on-ground action?*
- *Would you describe yourself as a pure or applied researcher?*

What are your most common sources of research funding?

Do you perceive them (the funding agencies) to be focussed more on applied or pure research?

How do you plan your research priorities?

- who influences this? Your own organisation? Industry? Other organisations? Other stakeholders? Funding agencies?

What are the key values/principles that guide your research?

Describe the nature of your research practice.....

- individual or team?
- one-off or programmatic?

EXTENSION OF RESEARCH

Who are the key stakeholders in relation to your research?

How does your research influence these stakeholders?

- *How do they get it?*
- *What impact has your past work had on practice?*
- *What impact is your current work likely to have?*
- *Whose responsibility is it to make the link between research and practice?*

If we look at the continuum, where along that is your own work focussed?

(overleaf)

How would you describe the linkages between research and extension?

- *Similar objectives?*

- *Mechanisms?*
- *Can they be improved? How?*
- *Risks*

Ideally, how would you picture research and extension working together?

UNDERSTANDING AND REPORTING ON OUTCOMES & IMPACT

What constitutes a successful research project for you?

How do you gather feedback on your research?

- *Who is important?*
- *What mechanisms provide you with feedback?*
- *Do you use this feedback to*
    - *prioritise/ change / focus what you do?*
    - *Improve the way you do your research – quality control, new techniques, etc*
    - *Improve the impact of your research?*
- *Learning / improvement?*

What would be the economic, environmental or social impact for a grower if they were to implement your research?

Have they thought about this?

Do you measure the rate of adoption of your research?

What do you think are the barriers or triggers to adoption?

Do you measure the impact of your research?

How?

Understand differences between adoption and impact?

Do you report on the outputs-outcomes and impacts of your research?

How?

What frameworks, if any does your organisation use?

Can the CRDC/CRC report formats be improved to include this more easily?

CRDC and other research groups are increasingly required to report on triple bottom line outcomes. Do you feel you have a role in this?

Why/why not?

How?

Whose role do you think this is?

What do you think would be helpful for TBL reporting and program evaluation? - resources, skills, help, support, frameworks, processes

Program evaluation focuses on the outcomes from whole research areas (eg IPM) rather than on outcomes of individual projects. What concerns, if any, would you have about program evaluation?

What motivates you to do research?

For you as an individual, are there any conflicts between personal, institutional and funding agency goals?

"Continuum"

Impact
Practice
Ownership, Understanding, Confidence
Integration and Knowledge Sharing
Knowledge Base – Management options
Knowledge Base – Fundamental
(Inputs)