

Part A

Proving our worth – A critique of our evaluation efforts

Grahame Collier
T Issues Consultancy

PO Box 728
Balgowlah NSW 2093

Phone: 612 9948 8243
Fax: 612 9948 8243
Mobile: 0419148551
Email: grahame@tissues.com.au
Web: www.tissues.com.au

Grahame Collier has been leading teams of people undertaking and evaluating strategic approaches to challenging social justice issues for most of his working life. He is now a sole trader consultant working primarily in public health and sustainability. His current work involves him in evaluating sustainability programs and projects.

Grahame is the Immediate Past President of the Australian Association for Environmental Education, a former member of the National Environmental Education Council and the NSW Council on AIDS, and a life member of Manly Warringah Life Line

Abstract: Proving our worth – A critique of our evaluation efforts

Few disciplines within the social change arena are required to ‘prove their worth’ more often and more rigorously than education. We are constantly expected to identify outcomes, chart progress, evaluate for behaviour shift and report on results. Other triggers for motivating change can get away with:

- *measuring volumes - for example, stormwater engineers can tell us the amount of green litter collected in a gross pollutant trap], or;*
- *counting numbers – for example, regulators can tell us how many people were fined for littering.*
- *offering hip pocket incentives – ‘this subsidy is available’ and counting uptake.*

For those of us in education though, we need to evaluate people and change. What it is that people are actually doing [and knowing and believing]; and are our programs having a direct effect on this? Counting the numbers of brochures distributed, or providing data about how many people attended our event is no longer good enough – and it never was!

This paper will investigate the issue of evaluation and it will critique our efforts to meet the demands that it places on us as worthy professionals in sustainability education. It will provide some case studies of quality evaluation programs and identify some of the common features of the best evaluation processes.

The phrase - ‘evaluation is challenging’ will not be used in this paper. Why not? Because proving our worth is crucial and undertaking a quality evaluation process is no less important, or difficult, than delivering a good education program.

Introduction

Undertaking evaluation of our efforts is ultimately satisfying; it requires intellectual rigour and clarity of purpose – it provides us with real information to show that we are doing something of worth and value.

Most of us as education professionals working in Education for Sustainability [EfS] are doing so because our hearts are in it. We want to know that our time is not being wasted our values are worthy and we are achieving what we set out to do.

Evaluation of our efforts is the only way to get some perspective about these issues – to prove our worth. Our need for effective evaluation is both personal and professional.

What do we mean by evaluation?

Definitions of evaluation abound. All contain similar concepts – judgement, worth, process and data or information. Two of the most succinct definitions are:

Evaluation is the process by which we judge the worth or value of something (Commonwealth Dept of Health and Aged Care 2001)

Evaluation is the systematic process of collecting credible data and using it to make judgements about the worth of a product, service or process at any point in the program's life cycle (NSW Department of Environment Climate Change and Water 2004).

It is essential to be clear about the concepts involved in effective evaluation at the outset of this paper. While evaluation involves judgement, the most meaningful evaluations base judgements on a range of objective data gleaned from a variety of sources and structured under a number of evaluation questions (NSW Department of Environment Climate Change and Water 2009). Many practitioners in sustainability education across Australia use the program logic model and its outcomes hierarchy approach to structure their evaluation processes (Funnell (1997) and Zammit, Cockfield and Funnell (2000)). This model involves the identification and organisation of outcomes in a hierarchy and the development of key evaluation questions for each outcome. Evidence is collected about the extent to which each question has been answered and hence: has the outcome been achieved (Department of Environment, Climate Change and Water. 2004).

Guba and Lincoln (1989) argue that evaluation identifies and raises understanding of local values and so the results cannot necessarily be generalised to other settings. The evaluation process engages people in a teaching/learning process, whereby evaluators, clients, sponsors, and stakeholders both teach and learn from one another. At its best it is a continuous process, because its 'findings' are continuously recycled and updated. They see evaluation as an emergent process; it cannot be fully designed in advance as its focus depends on inputs from stakeholders and its activities may need to vary through the process. Quality evaluation is about sharing accountability rather than assigning it.

The *Guide to using Social Research in Sustainability Programs (2009)* confirms that evaluation/research is useful, even in time constrained situations. This guide states that evaluation achieves a number of real benefits. It: starts discussions; creates a strong evidence base; improves programs and their outcomes; contributes to a body of knowledge, informs policy development to drive behaviour change; and involves practitioners in all steps of the process and in modifying their own practice.

Evaluating Behaviour Change

Because much of the work undertaken in EfS is about motivating shifts in people's behaviours (World Commission on Environment and Development 1987) this paper briefly discusses behaviour change before critiquing our evaluation efforts in the context of behaviour change more generally.

Social determinants: It is important to acknowledge at the outset that people's behaviour is established within a number of social determinants. Education and other behaviour shift motivating tools or levers are but a part of the mix of ways that can be used to influence behaviour. Factors including a person's upbringing, his or her socio-economic status, financial capacity, education level and capacity to act on available choices, form some of the context in which behaviour occurs. In health promotion theory, these are called the social determinants of health. (Kerr et al 2005). These social determinants also establish the context within which environmental sustainability change can be motivated. Evaluation of behaviour change should take account of the social determinants.

Levers for behaviour change: There are a significant number of levers apart from education that are used by government and other agencies to influence behaviour choices (Department of Environment Climate Change and Water 2006). These include:

- Economic stimulants [*Are there financial reasons why ... ?*] The economic stimulants might be positive or negative...incentives and/or penalties
- Regulatory stimulants [*Do I have to? What do the laws say?*]
- Peer group stimulants or social norms [*What is the normative behaviour in this society/community? I will do what they do.*]
- Public policy stimulants [*Are there water restrictions in place?*] These prompts might be seen as a combination of regulatory and social norms.
- Social structures [*Is society and its institutions supportive of the behaviour?*]. For example do council ordinances allow me to install a water tank at home?
- Opportunity [*Are real behaviour choices available now?*]
- Structural stimulants [*Are the structures in place to support this behaviour?*] For example, where would recycling be without kerbside collection programs?

A reasonable argument can be made that all of these levers should be evaluated for their impact on behaviour. There is a significant gap in our understanding about the specific outcomes that they achieve. In this article though, our focus is on education and proving that it works

Influencing behaviour: Collier and Smith (2009) identify that programs are successful at influencing behaviour when they:

- a. Consider evaluation an integral part of the design and delivery of the program.
- b. Use evaluative methods that demonstrate the results of the evaluation and indicate increases in knowledge, improvements in skills and/or changes in attitudes - and yield data on the intended or actual behaviour changes as a result of the program.
- c. Acknowledge that the process of evaluation involves the target/community in the review of the program and its performance.
- d. Demonstrate that the evaluation results feed into future program adaptation and development and encourage life-long learning.

Behaviours relevant to sustainable societies can be grouped into five key categories (Munroe 2003):

- Environmental activism (e.g., actively participating in or leading environmental initiatives).
- Non-activist political behaviours (e.g., joining an organisation, voting, signing a petition, or writing a cheque).
- Consumer & householder behaviours (e.g., purchasing 'green' products, recycling, reducing energy use, and altering consumption habits).
- Ecosystem behaviours (e.g., putting up bird boxes, planting sea oats, counting wildlife populations, promoting prescribed fire).
- Other behaviours which are specific to our expertise or workplace (e.g., reducing waste in the production process, establishing mortgage criteria for energy efficient houses, suing a polluter, etc.).

Education and change programs need to be clear about which behaviours and actions are being targeted before they begin to plan an effective evaluation process.

What is being changed and measured? The important part of evaluating behaviour change is to know, to be clear about; and report upon exactly what change is being measured. Is it behavioural intent, self reported behaviour and/or actual observable behaviour?

Behavioural intent

Much of what is relatively simple to evaluate is the behavioural intent of the target population as a result of our EfS efforts. Participants will be able to respond to questions about what they intend to do immediately after a workshop, as part of the evaluation of a social marketing campaign and/or at an event. Fishbein and Ajzen (1975) through their Theory on Reasoned Action, provide significant conceptualisation about the notion of behavioural intent. They argue, and many more recent publications support this position, that people need to have the intent to change before an actual change occurs. Gladwell (2000) calls this the tipping point, and stresses that this is the point at which a person or organisation is ready to make a change. Fishbein and Ajzen's original model sees norms and positive attitudes as forerunners to the intention to act, see below in fig 1.

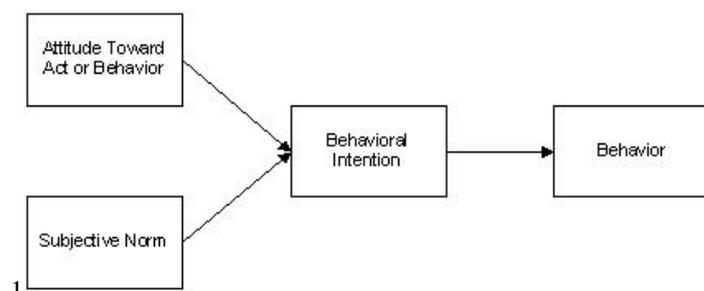


Figure 1: The Theory of Reasoned Action (TRA) (Source: Fishbein & Ajzen (1975))

In terms of evaluation, behavioural intent can be measured through a range of post intervention data collection processes all focusing on the key question: *What do you intend to do as a result of this program?*

Self reported behaviour

The second aspect of behaviour that can be measured is self reported behaviour change. While some people are concerned that self reporting can equal fabrication, the literature would dispute this fact as long as data is collected rigorously, Kerr (2005) and Owen (1999). Measuring self reported change is extremely important to proving our worth. The key aspects of the process include Wadsworth (1997) and American Evaluation Association (2004):

- Collecting data after the education program so that people have time to make the behaviour changes that they intended.
- Carefully triangulating data collection processes and questions so that data can be checked. For example in a questionnaire about installation of water tanks, a question might be asked in two or more ways: *When did you install your new water tank? What is the best thing about having a water tank? Did you apply for the subsidy when you got your new tank? What do your neighbours say about you having a water tank?* A review of the consistency between the answers to these questions will give some certainty about whether the change has actually occurred.

- Acknowledge that at times self reported behaviour is the best form of information that you can get. How else are you going to measure if people across a whole community are having shorter showers?

Actual quantifiable/observable behaviour

In some programs, it is possible to get information that provides real quantifiable and/or observable evidence about behaviour that has changed as a result of a program. Programs where this information is collected and made a part of the evaluation report are very robust but also quite resource intensive. Over time auditing programs, the use of energy and/or water metres, review of energy and/or water bills, observation programs: Curnow (2003) and industry programs that provide input and output data, are examples of programs where actual behavior/operational changes can be measured. Most of the time and within constrained resources, this is just not possible.

Critiquing our Evaluation Efforts

In EfS we have certainly moved a long way forward over the last decade in our focus on quality evaluation processes; guides are being produced, professional development programs are available, case studies can be found on-line, conference presentations abound. Funding authorities are demanding evaluated projects. But there is more to do. Much of the evaluation that is occurring does not find its way into the literature, so the suggestions that follow are based primarily upon the personal perspectives and experiences of the author.

We evaluate behaviour change relatively poorly and can improve this easily

In part this is because we are not clear about what we are measuring [see above]. Are we measuring: behavioural intent/self reported change/quantifiable/observable change? High quality program design will make it clear what can be measured because the outcomes of the project will be oriented towards influencing behavioural intent or actual behaviour. If the outcome says actual behaviour and only intended behaviour is measured then there is a problem.

In part we evaluate behaviour shift poorly because we expect change to occur quickly and we don't measure for it over a long enough period of time. Good project design will identify how long we expect it to be before change occurs. The timing of the evaluation must mirror the timing of the expectation for change.

Also we evaluate behaviour change poorly because we provide insufficient resources to the process of evaluation – to collect the available data. Hence we are unable to follow up sufficiently even to measure for self reported change [let alone actual changes]. Evaluation costs money and if we want to prove our worth, up to 10% of the project's budget needs to be put towards an effective evaluation.

We begin planning for evaluation too late in the process

While evaluation can begin at almost any point in an initiative, ideally you should plan, agree to, and budget for evaluation from the outset. It should become part of your day-to-day management practice and support the processes involved with developing and implementing a promotion and prevention initiative. Commonwealth Dept of Health and Aged Care (2001).

Evaluation must be considered at every stage of program planning and implementation to ensure that adaptive management occurs as part of continuous improvement. The more integrated program planning and evaluation planning, the better the quality of the evaluation process. Often we leave the evaluation planning until it is too late to collect even the most basic of data.

We develop too many outcomes and make evaluation too dense

Although the move to the development of the program logic approach is encouraging, Department of Environment, Climate Change and Water. (2004) there is a tendency to develop too many outcomes for a project. If this occurs it is difficult to reflect these within an outcomes hierarchy and to collect data about the extent to which each one has been achieved. Program designers need to be rigorous in identifying outcomes that are ‘immediate’, ‘intermediate’ and ‘ultimate’ and that collect key concepts together. It is the author’s experience that projects with no more than ten outcomes are more effectively evaluated than those with more outcomes.

We are unclear about what we are evaluating

The American Evaluation Association (2004) states that in order to be effective, evaluation:

- *assesses the impact, appropriateness, effectiveness, efficiency and legacy of policies and programs and a process to promote accountability to the community, funding bodies and Council.*
- *is essential to the professional credibility of program managers in that if they are developing and implementing a program, it is important to know that it is working.*
- *is best scoped within and overarching conceptual framework for evaluating programs with an emphasis on learning, improvement and accountability.*
- *guides the development and implementation of program-level and investment-level plans.*
- *uses data that is relatively easily collectable and reasonably accessible to all sectors of the community, under conditions that do not inhibit its use.*

In each evaluation it is essential to be clear about what is being evaluated.

Evaluation does not guide our future practice well enough

Often we forget the formative nature of the evaluation that we are undertaking. In part this is because evaluation often occurs toward the end of a time limited, grant funded project, where the immediate task is getting the project finished. In these circumstances it is not surprising that the reflective processes that should accompany every evaluation are forgotten. In every circumstance it is important to fully consider the key question – *What did I learn from that project for the future?* A rigorous evaluation process will guide thinking about this question.

We have not sorted the principles that underpin our evaluation efforts sufficiently

The United Nations Development Program (2006) states that evaluation should be:

Independent—*Management must not impose restrictions on the scope, content, comments and recommendations of evaluation reports. Evaluators must be free of conflict of interest.*

Intentional—*The rationale for an evaluation and the decisions to be based on it should be clear from the outset.*

Transparent—*Meaningful consultation with stakeholders is essential for the credibility and utility of the evaluation.*

Ethical—Evaluation should not reflect personal or sectoral interests. Evaluators must have professional integrity, respect the rights of institutions and individuals to provide information in confidence, and be sensitive to the beliefs and customs of local social and cultural environments.

Impartial—Removing bias and maximizing objectivity are critical for the credibility of the evaluation and its contribution to knowledge.

Of high quality—All evaluations should meet minimum quality standards

Timely—Evaluations must be designed and completed in a timely fashion so as to ensure the usefulness of the findings and recommendations

Used—Evaluation is a management discipline that seeks to provide information to be used for evidence-based decision making. To enhance the usefulness of the findings and recommendations, key stakeholders should be engaged in various ways in the conduct of the evaluation.

For all evaluation of EfS programs, these principles or some just like them need to be considered and adopted or adapted. Often in our attempts to evaluate on the run, we forget that there are solid principles that underpin all evaluation.

Some Leading Practice Case Studies of Evaluation

Finding case studies of good quality evaluation is somewhat difficult, not because they do not exist, but because evaluation reports often stay in-house or are only sent to funding providers, They are not reported in the literature.

Living Smart Homes (LSH) is designed to engage homeowners with sustainable practices through a range of interactions. Participants were asked to change household behaviours and to discuss the changes and the barriers to participation in the program and to making the behavioural changes. The program influenced participants' actions, with over half (56%) believing that the information they learnt in the 'Living Smart Homes' program would influence their behaviour forever. In a self-assessment about the impact of the program approximately half reported making at least some behavioural changes in the areas of Energy (67%), Waste (61%), Water (48%) and Transport (48%). Calculations from the LSH website modules estimate the annual collective savings of Living Smart participants, if they were to continue their behaviour for a year, to be 224,350 kilograms of greenhouse gases, 3,740 kilolitres of water, 308,963 kilometres of travel, 150,624 kilowatt hours of energy and 37.48 tonnes of waste not deposited in landfill. Miller W (2009)

At the local level baseline data about behaviour change is being gathered and compared to information for other sources For example at Warringah Council, evaluation of their community workshops series on sustainability contains questions drawn from the *Who cares about the environment in 2009* study - Department of Environment Climate Change and Water (2009). This allows for comparison of data from local responses against state wide benchmarks. It also enables comparison over time given that state-wide data is available from 1994 onwards.

Pillora (2010) provides a case study about a Blue Mountains City Council project. Council undertook an evaluation using the outcomes hierarchy model in working with three community education projects in the local government area. The approach reviewed the use of a method of engagement identified as appreciative enquiry and a rigorous evaluation process was undertaken of the effectiveness of this method. This publication also provides an

evaluation case study of work undertaken by Wollongong City, Kiama and Shellharbour Councils.

The NSW Department of Environment and Climate Change (2008) publication, *Educating the Community About Litter* [www.environment.nsw.gov.au/education/littered.htm] provides a case study of the evaluation of a large social marketing litter reduction program.

The NSW Climate Consensus Project, an initiative of the Nature Conservation Council of NSW (NCC), utilised deliberative democracy to enable representative groups of everyday citizens in NSW to contribute to climate change policy and program development at both a local and State government level. Through a series of 12 local forums and a state-wide NSW Community Climate Summit and in partnership with 14 local councils, randomly selected citizens, who represented the wider community demographic, deliberated over the question *how can we work together to respond to climate change*. A comprehensive approach to evaluation was used including how feedback received can be better broken down into six key sources: council staff and community educators involved in the training program; local forum participants, local council partners; summit participants; summit facilitators and the NSW Government. Feedback from each Forum contributed to a significant evaluation report. For information contact the NCC.

Conclusion

We evaluate for personal and professional gratification, to prove our worth and the worth of education as a lever for change. Rigorous evaluation of EfS programs will continue to be important for the future and this requires educators to understand the evaluation process and to work cleverly to ensure that good evaluation of education for change is undertaken. We need to have an understanding of the theory of evaluation and the capacity to put this into practice so that we can legitimately prove our worth.

References

American Evaluation Association (2004) *Guiding Principles for Evaluators* see: www.eval.org.

Commonwealth Dept of Health and Aged Care (2001) *Evaluation: A Guide for Good Practice*

Collier G. and Smith P (2009) *Beyond Lip Service: A Council Approach to Planning for Behaviour Change* *Australian Journal of Environmental Education*, vol. 25.

Curnow R, Spehr K & Casey D. (2003) *Keeping It Clean – Changing Disposal Behaviour*. Unpublished see: <http://www.communitychange.com.au/downloads/KeepingItClean.pdf>

Department of Environment Climate Change and Water (2009) *Who cares about the environment in 2009*.

Department of Environment Climate Change and Water (2006). *Learning for Sustainability: NSW Environmental Education Plan 2007-10*

Department of Environment, Climate Change and Water; (2009) *Guide to using Social Research in Sustainability Programs*.

Department of Environment, Climate Change and Water. (2004). *Does Your Project Make a Difference*

Department of Environment and Climate Change (2008) *Educating the Community About Litter 2000 to 2003*

Department of Environment Water, Heritage and the Arts (2009). *Living Sustainably. The Australian Government's National Action Plan for Education for Sustainability*.

Funnell S. (1997): *Program Logic: An Adaptable Tool for Designing and Evaluating Programs* in *Evaluation News and Comment*, Vol 6, number 1, July 1997, Australasian Evaluation Society, ACT.

Fishbein, M. & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behaviour: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.

Gladwell M. (2000) *The Tipping Point* http://wikisummaries.org/The_Tipping_Point

Guba E. and Lincoln Y. (1989) *Fourth Generation Evaluation* Sage: Thousand Oaks Publishing.

Kerr J, Weitkunat R. and Moretti M. eds [2005] *ABC of Behaviour Change: A guide to successful disease prevention and health promotion*. Elsevier

Miller E. Buys L. and Bell L. (2009) *Living smart homes: a pilot Australian sustainability education program*. *Journal of Education for Sustainable Development*.

Monroe M. (2003) *Two Avenues for Encouraging Conservation Behaviours*, *Human Ecology Review* Vol.10 No.2 2003

Owen J. with Rogers P. (1999) *Program Evaluation: Forms and Approaches* 2nd Edition, Allen and Unwin,

Pillora S. and Blackburn N. (2009) *Sustainability Tool Selector: A Guide for Local Government*. Prepared for the Urban Sustainability Support Alliance for the Institute of Sustainable Futures. University of Technology, Sydney

World Commission on Environment and Development (1987) *Our Common Future*. Report of the World Commission on Environment and Development, United Nations.

United Nations Development Program (2006). *The Evaluation Policy of UNDP*, Available at: <http://www.undp.org/eo/documents/Evaluation-Policy.pdf>

Wadsworth Y (1997). *Everyday Evaluation on the Run*. Allen & Unwin, 2nd Edition.

Zammit C. Cockfield G. and Funnell S. (2000) An outcomes-based framework for evaluating natural resources management, policies and programs, Published by Land and Water Australia Social & Institutional Research Program