HEALTH PROMOTION INTERNATIONAL
Vol. 13, No. 1

Printed in Great Britain

Evaluating health promotion—progress, problems and solutions[1]

DON NUTBEAM
WHO Collaborating Centre for Health Promotion, University of Sydney, Sydney, Australia

SUMMARY

Several issues of current debate in health promotion evaluation are examined. These include the definition and measurement of relevant outcomes to health promotion, and the use of evaluation methodologies which assess both the outcome achieved and the process by which it is achieved. Considerable progress is being made in understanding the complexity of health promotion activity, and in the corresponding need for sophisticated measures and evaluation research designs which reflect this complexity. The more powerful forms of health promotion action are those which are long term, and least easily predicted, controlled and measured by conventional means. Against this, important and valued advances in knowledge and credibility have come from more tightly defined and controlled interventions, which have been evaluated through the application of more traditional experimental designs. This tension between 'scientific rigour' and the perceived advantages (in long-term effectiveness and maintenance) coming from the less-well-defined content and methods of community controlled programmes continues to pose technical problems in evaluation. It is important to foster and develop evaluation designs which combine the advantages of different research methodologies, quantitative with qualitative, in ways which are relevant to the stage of development of a programme. The use of a diverse range of data and information sources will generally provide more illuminating, relevant and sensitive evidence of effects than a single 'definitive' study. Evaluations have to be tailored to suit the activity and circumstances of individual programmes—no single methodology is right for all programmes.

Key words: evaluation; health promotion; measurement

INTRODUCTION

In past 20 years there has been an enormous growth in the volume of research which is of relevance to health promotion. This rapidly expanding research base has advanced knowledge and improved understanding of the determinants of health in populations, and how to bring about change in those determinants to improve health.

The evolution of the concept of health promotion, especially in the decade since publication of the *Ottawa Charter for Health Promotion* [World Health Organization (WHO), 1986], has added sophistication to this analysis, greatly expanded the range of strategies and actions to promote health, and in doing so, greatly complicated the challenges of evaluating health promotion.

There are many different interpretations of what represents 'value' from a health promotion programme. Among the perspectives reflected in the literature is that of the population who are to benefit from health promotion action who may place great value on the ways in which a programme is conducted, particularly whether or not the programme is participatory, and addresses priorities which the community itself has identified; that of health promotion practitioners who

[1] Edited text of a paper prepared for the Fourth International Conference on Health Promotion, *New Players for a New Era—Leading Health Promotion into the 21st Century*, Jakarta, 21–25 July 1997.

need to be able to judge with reasonable confidence the success of a programme in relation to its defined objectives, as a form of feedback on performance; that of managers who need to be able to judge the success (or likely success) of programmes in order to make decisions about how to allocate resources, and be accountable for those decisions to funders, including the community and elected representatives; and that of academics who need to be able to judge success in order to improve knowledge and understanding of cause and effect in interventions (Green, 1987; Weir, 1991; Holman *et al.*, 1993; Viney, 1996).

Correspondingly, there is currently a vast spectrum of approaches to evaluation which are used in health promotion. These range from highly structured, methodology-driven evaluations, exemplified by randomised controlled trials, through to much less rigidly structured, highly participative forms of research and evaluation. Making sense of this diversity has been a challenge, and several structured models for planning and evaluation of health promotion programmes have been developed as a part of the response to this challenge (Sanderson *et al.*, 1996).

This emphasis on structure and sequence in health promotion planning has been important in establishing the credibility of health promotion as a form of public health action, and as a distinct discipline in the health sciences. Such developments are reflected in the substantial growth in the number of textbooks on health education and health promotion, including those specifically directed towards evaluation (Windsor *et al.*, 1984; Green and Lewis, 1986; Hawe *et al.*, 1990). This 'development' may also have had unintended consequences by narrowing the definition of what constitutes health promotion and the criteria for its evaluation—a dilemma referred to below.

Debate continues about what represents 'good practice' in research and evaluation in health promotion (Nutbeam, 1996a). This paper provides an overview of progress in evaluation, and identifies some contemporary dilemmas before concluding with discussion of how further progress may be made.

DEFINING 'SUCCESS' IN HEALTH PROMOTION

Valued outcomes and valued processes in health promotion

At its core, evaluation concerns assessment of the extent to which an action achieves a *valued* outcome. In most cases there is also value placed on the process by which these outcomes are achieved. The *Ottawa Charter* identifies both valued outcomes and valued processes in health promotion. In the *Charter*, health promotion is defined as follows:

> health promotion is the process of enabling people to exert control over the determinants of health and thereby improve their health.

Health promotion is described as a 'process', indicating that it is a means to an end, and not an outcome in its own right. Health promotion is an activity directed towards enabling people to take action. Thus, health promotion is not something that is done *on or to* people, it is done *with* people, either as individuals or as groups. Participation and partnership are valued processes in health promotion.

The purpose of this activity is to strengthen the skills and capabilities of individuals to take action, and the capacity of groups or communities to act collectively to exert control over the determinants of health. Thus, empowerment of individuals and communities are valued outcomes.

In tackling the determinants of health, health promotion will include both actions directed towards changing determinants within the more immediate control of individuals, including individual health behaviours, and those factors largely outside the control of individuals, including social, economic and environmental conditions. Thus, actions which support healthy lifestyles and create supportive environments for health are also valued outcomes to health promotion.

Poor definition and measurement of anticipated outcomes to health promotion activities has long been considered a stumbling block to progress (Green and Lewis, 1986, pp. 5–6). Better definition of anticipated outcomes precedes more relevant and convincing evaluations of health promotion programmes and activities, and better communication of what constitutes 'success' in health promotion.

Health outcomes and definitions of 'success' in health promotion

In many countries at the current time, considerable attention is being given to health outcomes (Epstein, 1990). A health outcome in such cases can be defined as a change in the health of an individual or group which is attributable to an intervention or series of interventions.

The driving force behind the attention to health outcomes is the perceived need to improve the effectiveness and efficiency of investments made by people and their governments in health, particularly in health services—so that the spotlight falls less on what is done and more on what is achieved (Sheill, 1997). It is argued that concentrating on outcomes (rather than on inputs in the form of medical procedures/hospital beds and so on) provides a more rational way of deciding on what interventions will achieve greatest health gain for a given investment.

At face value, health promotion would do well in such an environment, offering the potential of substantial health gains for relatively modest investment relative to other forms of health service. However, the rather complex and distant relationship between typical health promotion activities and 'health outcomes' expressed in terms of change in physical function or disease state, combined with the paucity of *evidence* relative to mainstream health system activities has made it difficult to achieve the progress which might be justified. Additionally, health outcomes which are defined mainly in terms of physical function or disease state, are not necessarily the same as the 'valued outcomes' from health promotion referred to above.

Health outcomes and health promotion outcomes

Given this context, it is important to distinguish between the different types of outcome associated with health promotion activity, and to articulate the relationship between health promotion outcomes and the type of health outcomes commonly referred to in the definition given above. In an effort to do this, different forms of health outcomes hierarchies and models have been developed to explain the relationship between health promotion action and health outcomes (Tones, 1992; King, 1996; Macdonald *et al.*, 1996).

Figure 1 shows a model which illustrates these relationships. Three different levels of outcome are identified (Nutbeam, 1996b).

Health and social outcomes

In the *Ottawa Charter*, health is defined as 'a resource for life, not the object of living'. The social outcomes reflect this functional definition of health and in the model represent the top of the hierarchy—the end point of health and medical interventions. Thus, outcomes such as quality of life, functional independence and equity have the highest value in this model. Related to this, though not the only influential factor, are health outcomes which are more narrowly defined in terms of disease experience, physical and mental health status.

Intermediate health outcomes

Intermediate health outcomes represent the determinants of health and social outcomes. Health promotion is directed towards increasing people's control over such determinants. Personal behaviours which provide protection from disease or injury (such as physical activity), or increase risk of ill-health (such as tobacco use) are represented through the goal of 'healthy lifestyles'. The physical environment can limit access to facilities, or represent a direct hazard to the physical safety of people; and economic and social conditions can limit people's participation in society. These determinants are represented as 'healthy environments'. These environments can both have an impact directly on health and social outcomes, and indirectly influence healthy lifestyles by making individual behaviours more or less attractive (for example, by limiting or enhancing access to facilities for physical activity). Access to and appropriate use of health services are acknowledged as an important determinant of health status and are represented as 'effective health services'.

Health promotion outcomes

Health promotion outcomes reflect modification to those personal, social and environmental factors which are a means to improving people's control and thereby changing the determinants of health (intermediate health outcomes). They also represent the more immediate results of planned health promotion activities.

The cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health,

are summarised as 'health literacy' in the model. Examples of health promotion outcomes would include improved health knowledge and motivation concerning healthy lifestyles, and knowledge of where to go and what to do to gain access to health and other support services. Furthermore, like literacy itself, health literacy means more than being able to read pamphlets and make appointments. In the same way that literacy can be empowering by giving people the necessary skills and confidence (self-efficacy) to participate in everyday activities, including the political process, so too is health literacy intended to reflect this larger concept.

'Social influence and action' includes organised efforts to promote or enhance the actions and control of social groups over the determinants of health. This includes mobilisation of human and material resources in social action to overcome structural barriers to health, to enhance social support, and to reinforce social norms conducive to health. Examples of outcomes would range from improved social 'connectedness' and social support, through to improved community competency and community empowerment.

Healthy environments are largely determined by 'healthy public policy and organisational practices'. Policy-determined legislation, funding, regulations and incentives significantly influence organisational practice. Thus examples of outcomes here would be changes to health and social policies directed towards improving access to services, social benefits and appropriate housing, and changes to organisational practices intended to create environments which are supportive to health.

Health promotion actions

Figure 1 also indicates three health promotion actions—what to do, as distinct from what outcomes are achieved. 'Education' consists primarily of the creation of opportunities for learning which are intended to improve personal health literacy, and thereby the capacity of individuals

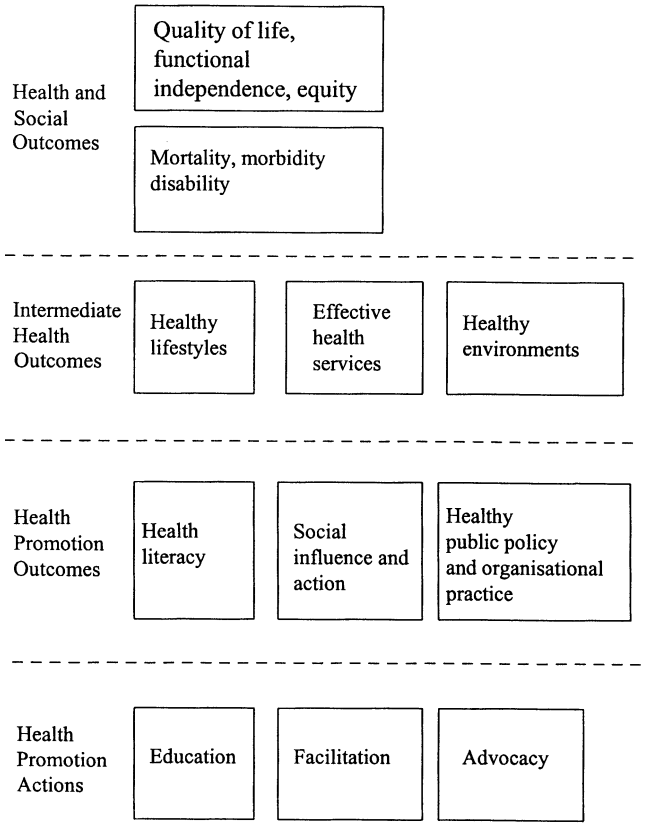

Fig. 1: An outcome model for health promotion.

and communities to act to improve and protect their health. 'Facilitation' is action taken in partnership with individuals or social groups to mobilise social and material resources for health. 'Advocacy' is action taken on behalf of individuals and/or communities to overcome structural barriers to the achievement of health.

The figure can be used to illustrate the linkages not only between the different levels of outcomes, but also *within* levels. For example, among the intermediate outcomes, action to create healthy environments may be both a direct determinant of social and health outcomes (for example, by producing a safe working and living environment, or improving equity in access to resources) and separately influence healthy lifestyles, for example by improving access to healthy food, or restricting access to tobacco products.

Implicit in the figure is the notion that change in the different levels of outcome will occur according to different time-scales, depending on the nature of the intervention and the type of social or health problem being addressed.

There is a dynamic relationship between these different outcomes and the three health promotion actions, rather than the static, linear relationship which might be indicated by the model in Figure 1. Health promotion action can be directed to achieve different health promotion outcomes by shifting the focus or emphasis to an intervention. Deciding on what represents the best starting point and how to combine the different actions to achieve valued health promotion outcomes through valued processes are at the core of 'best practice' in health promotion.

Measurement of outcomes in health promotion

The definition and measurement of outcome indicators which are relevant to the intervention, and sufficiently sensitive to reflect intervention effects has been a long-standing challenge in health promotion. The health promotion outcomes indicated in Figure 1 are closest to the health promotion action, and thereby the most sensitive 'object of interest'—the most likely to be heard beyond the background noise of everyday community activities. These health promotion outcomes will, in turn, be directed towards intermediate health outcomes (health behaviours, healthy environments and effective health services), these may be the most relevant 'objects of interest', as they represent more widely understood and accepted outcomes to health promotion activity. Measurement of change in these intermediate health outcomes using appropriate indicators may also be necessary, depending on the size, comprehensiveness and duration of the intervention.

In the past, greatest attention has been given to the development of valid and reliable methods for the measurement of health behaviours and, to a lesser extent, measurement of indicators of the physical environment, and changes in patterns of health service provision and utilisation. Greater attention needs to be given to the development of measures which are more sensitive to the immediate impact of health promotion actions.

The range of strategies employed in health promotion and different outcomes from those strategies has meant that a wide range of potential indicators need to be considered. Whilst it is not possible to provide an exhaustive list here, examples of indicators which can be used in assessing the achievement of health promotion outcomes include the following.

- **Health literacy:**
  knowledge relevant to the problem of interest;
  self-confidence in relation to defined tasks (self-efficacy);
  self-empowerment;
  attitudes and behavioural intentions;
  future orientation;
  participation in health promotion programmes.
- **Social mobilisation:**
  community competency;
  community empowerment;
  social capital;
  social connectedness;
  peer and community norms;
  public opinion and public mandate for policy action;
  community ownership of health promotion programmes.
- **Public policy and organisational practice:**
  policy statements;
  legislation and regulations;
  organisational procedures, rules and administrative structures;
  management practices;
  funding and resource allocation;
  institutionalisation of health promotion programmes.

Acknowledgement and adoption of such a range of measures of success fits more comfortably with modern concepts of health promotion. It would do much to move evaluation on from a

reductionist, psycho-social and behavioural paradigm in the assessment of success, to a more 'expansionist' approach to thinking about health promotion and the meaning of 'success' in health promotion.

One important block to this move is a paucity of reliable and valid measures of many of the indicators of success identified above. The definition and measurement of intermediate health outcomes such as health behaviours and healthy environments, and the health promotion outcomes which may influence them, has taxed the skills of researchers for decades. The task may be relatively straightforward in the case of defining and measuring smoking behaviour using information provided by individuals, but more complex in other areas such as assessing dietary behaviour or patterns of physical activity. Measuring knowledge, attitudes or values, personal and social skills as indicators of health literacy, community ownership of programmes and community empowerment as measures of social mobilisation, and organisational practice and public policy are potentially even more problematical.

The solution to many of these problems has rested in the construction of questionnaires, tests, scales and interview protocols. Such research tools are not only used to obtain information from individuals on personal knowledge, attitudes and behaviours, but can also be used to obtain information from relevant respondents on organisational policy and practice, and on community capacity and competence. Although there are no comprehensive 'tool kits' for outcome measurement in health promotion, much has been learned through careful experimentation in the past decades.

Currently, greater attention is being given to the development of indicators and instruments which measure changes in the health promotion outcomes listed above. Progress in this arena has been supported through a number of WHO publications (Abelin *et al.*, 1987; Badura and Kickbusch, 1991), as well as editions of specialist journals (Noack, 1988; De Vellis *et al.*, 1995). Much work remains to be done to develop sensitive, reliable and valid indicators for health promotion and intermediate health outcomes outside of the established comfort zones of the measurement of health behaviour and its psycho-social determinants, but progress is being made through experimentation in some of these more difficult areas (Kar *et al.*, 1988; Cheadle, 1992). More specific efforts include those directed at assessing individual and community empowerment (Schultz *et al.*, 1995), community action for health (Paine and Fawcett, 1993), policy advocacy (Schwartz *et al.*, 1995), and organisational change (Goodman *et al.*, 1993). Where available, more consistent use of established indicators and measurement techniques would not only do much to improve confidence in standards, but also have the additional benefit of increasing comparability between studies.

HOW BEST TO EVALUATE 'SUCCESS'? STAGES OF RESEARCH AND EVALUATION

A hierarchy of evaluation stages

Research to support the development of different health promotion interventions takes many forms. The model provided in Figure 2 is developed from an earlier version by the author and colleagues. It indicates six stages of research which go together to develop and evaluate a health promotion intervention (Nutbeam *et al.*, 1990). These include the following.

Stage 1: problem definition

This stage draws upon basic epidemiological research to investigate the causal basis and scope for an intervention, and community needs assessment to identify community concerns and priorities, to identify access points to reach and work with key individuals and populations, and to enable more direct community participation in problem definition and solution generation. This information defines the major health problems experienced within a defined population, the determinants of those problems, and the scope for change in those determinants.

Stage 2: solution generation

This stage draws upon social and behavioural research to improve understanding of target populations, and the range of personal, social, environmental and organisational characteristics which may be modifiable to form the basis for intervention, and intervention theory development can help to explain and predict change in individuals, social groups, organisations and the political process. Such theories and models are particularly useful in identifying plausible methods for achieving change in the personal, social and environmental characteristics referred

to above, and the potential for general application in different settings and with different population groups. This information clarifies the potential content and methods for intervention, and further defines the different needs of populations.

Together, stages 1 and 2 describe the cause, content, population and method which form the basic building blocks for planning health promotion interventions. Such information will describe a problem, can identify determinants of that problem, can indicate individuals, groups, institutions and policies in a defined community which are most in need of attention, and through this analysis, propose likely solutions. These possibilities can be narrowed and defined in terms of programme objectives which state the expected health promotion outcomes from a planned action (Hawe *et al.*, 1990, Chapter 3). Once such programme objectives have been defined, evaluation of a programme becomes more feasible. These programme objectives are the immediate, short-term focus for evaluation. Related intermediate health outcomes, and health and social outcomes may also be described at this time.

Finding a successful and sustainable solution to a defined health problem requires the systematic development and testing of an intervention. A staged approach to the development and testing of innovations has been recommended by several different authors (Flay, 1986; Sanson-Fisher *et al.*; 1996; Oldenburg *et al.*, 1996). Figure 2 describes a staged approach to evaluation research, indicating how the two fundamental tasks in evaluation research of *assessing outcomes* in order to determine the extent to which the intervention achieved what it was established to achieve, and *understanding the process* in order to identify the basic conditions for successful implementation of an intervention, and allow for reproduction of the intervention and subsequent repetition of successful outcomes.

Stage 3: testing innovation

Ideally, in order to establish evidence of success, evaluation of a new programme will go through these different stages. The relative importance of the two evaluation tasks will vary as an intervention goes through different stages of development. The figure indicates a hierarchy of study beginning with experimental studies which concentrate primarily on the question of whether or not an intervention achieves its desired outcomes. The function of such studies is to assess the extent

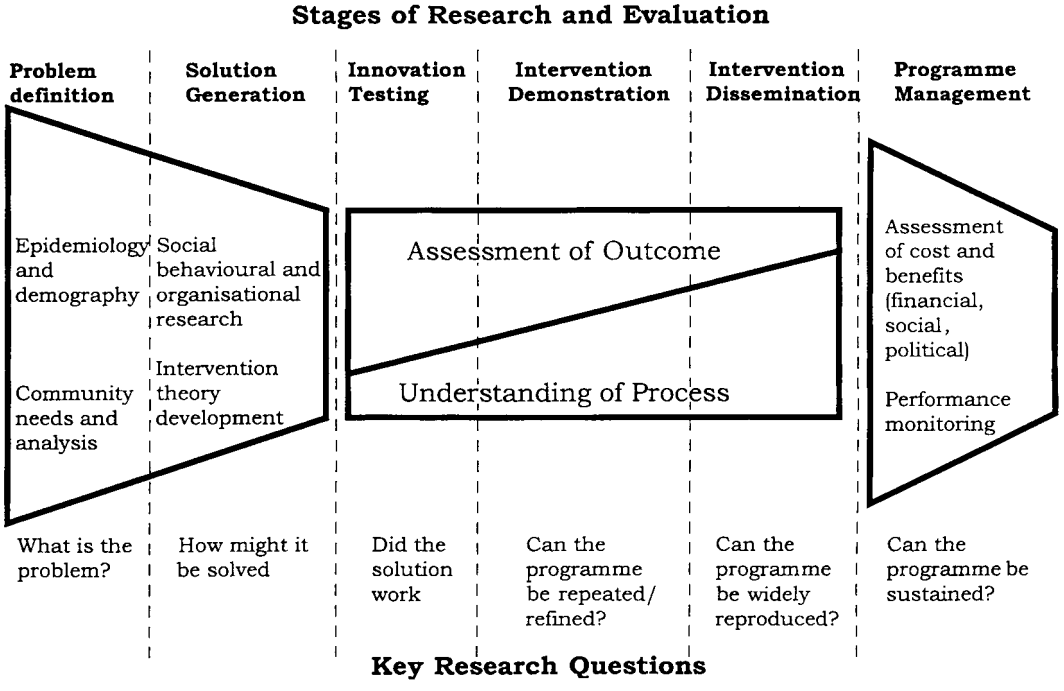

Fig. 2: Six-stage development model for the evaluation of health promotion programmes.

to which defined objectives can be achieved operating in the best possible conditions for success.

Because such studies need to be developed in such a way as to meet tightly defined standards, they tend to be of greatest interest to academic researchers. However, for the same reasons, such studies are often developed using resources and methods which are not easily reproduced, nor do such studies invite active participation in decision-making by the individuals and communities they are intended to benefit—a valued process in health promotion.

Stage 4: intervention demonstration

The forth stage, demonstration studies, shows a shift in the relative emphasis given to assessing outcomes and understanding process. If an intervention achieves the desired outcomes under ideal circumstances, the emphasis of the evaluation changes to consider more closely identification of the conditions for success. Here the task is to reproduce a programme in circumstances which are closer to 'real-life' and which better reflect the valued processes in health promotion, including control in decision-making, and the development of capacity for sustaining effects. This stage helps to clarify whether or not the desired outcomes can be achieved in a less artificial environment, and represent a reasonable investment of resources.

Such studies are of greater relevance and interest to communities and their leaders, as well as health promotion practitioners and activists, as they indicate that desired outcomes may be achievable in circumstances closer to real life. Specifically, they take account of the contextual variables of health promotion practice, and indicate the essential conditions which need to be established. Because of the balanced emphasis on both process and outcome, this type of study often produces more practical guidance, for example by indicating the importance of building community competency and working across sectors, as well as clarifying the resources which need to be committed for success. This stage in the process offers the opportunity for assessment of costs and benefits more related to real life conditions (Cohen, 1994).

Many programmes operating at community level would fall into this category. Practitioners and activists identify new ideas and programme strategies through the literature and/or word of mouth and seek to modify them to local circumstances. Although there are a growing number of studies of this type appearing in the research literature, such evaluation research appears to be less attractive to academic researchers who may be less comfortable with the uncertainties and lack of control over methodology and intervention in such studies, as well as the attendant reduction in the chances of demonstrating an intervention effect.

Stage 5: intervention dissemination

The fifth stage, dissemination studies, indicates a shift in emphasis still further. Here, attention is given to identifying the ways in which successful programmes can be widely disseminated. Such studies include those directed at improving understanding of the ways in which communities can be supported to adopt and maintain innovations (Jackson *et al.*, 1994; Rissel *et al.*, 1995) and build capacity (Hawe *et al.*, 1997), as well as studies of communities and organisations to determine how best to create the necessary conditions for success in different settings (Orlandi, 1986; Goodman and Steckler, 1987; Parcel *et al.*, 1989; Allensworth, 1994).

This type of evaluation research also provides information of great interest to communities, managers and practitioners because it helps to define what needs to be done, by whom, to what standard, and at what cost. This type of research is least common in the health promotion research literature, partly reflecting a lack of interest (and reward) on the part of academic researchers, and partly as a natural consequence of decline in the number of interventions which reach this stage of development (i.e. of proven efficacy) (Rychetnik *et al.*, 1997).

Stage 6: programme management

Beyond this stage, the basic evaluation tasks are directed towards supporting programme management. These tasks include monitoring the quality of programme delivery relative to the optimal conditions for success, and assessing value for money. The assessment of 'quality' in health promotion has been given considerable attention in the recent past, and a number of guides and manuals have been produced to assist with this task (Catford, 1993; Coppel *et al.*, 1994; Van Driel and Keijsers, 1997). The long-term management of programmes is not considered in detail in this paper.

The relative importance of the two major dimensions to evaluation research (outcome and process) will vary with a project's stage of devel-

opment and the target audience for the evaluation. Figure 2 indicates a hierarchical model, setting out the principal evaluation question at different key stages of development, and illustrating how the balance of importance between outcome and process evaluation changes at each of the three central stages.

EVALUATION OF OUTCOME: ASSESSING CAUSE AND EFFECT

By linking Figures 1 and 2, it should be obvious that it is hard to identify a simple causal chain which links a health promotion action to changes in health status. Such a simplistic 'reductionist' model for health promotion and disease prevention has long been discredited (Syme, 1996). The link between health promotion action and eventual health outcomes is usually complex and difficult to trace—a fact which poses real dilemmas in evaluations which seek to use social and health outcomes as primary measures of 'success'.

For example, smoking is a major cause of illness and disability which threatens the quality of life of many people. Quitting smoking or never starting will greatly reduce the future incidence and prevalence of several major causes of premature death, disease and disability. But even here, where the link between a behaviour and health outcome is clearly established, the relationship between different forms of health promotion intervention—education, behavioural counselling, changing social attitudes, environmental restrictions and price increases—and subsequent decisions by an individual to quit or not to start, are very complex (Chapman, 1993). Where the relationship is less well established or acknowledged—for example the relationship between income distribution or employment status and health (Kaplan *et al.*, 1996)—defining a causal chain between actions designed to alleviate the health impact of these determinants, and subsequent health outcomes becomes even more problematic. Currently, far more attention is being given to the complexities of these relationships, and the implications for public health action to respond to them (Evans *et al.*, 1994).

Given this situation, great attention needs to be given to clarity in the definition of health promotion outcomes, and to the evidence which indicates their relation to intermediate health outcomes, and subsequent health and social outcomes. Based on this model, evaluation of health promotion action should be based on measurement of change in the three types of health promotion outcome—achievement of improved personal health literacy, changes to public policies and organisational practices, and changes to social norms and community actions which, individually or in combination, increase people's control over the determinants of health.

In assessing the outcome to an intervention, two basic questions have to be addressed, namely:

(i) can change be observed in the object of interest; and

(ii) can this observed change be attributed to the intervention?

In this paper it is not possible to discuss in detail the full range of methodological issues that can arise in developing and executing an evaluation design in answering these questions. Some of these issues, such as sample size and selection, data collection techniques, and response rates, are common to all forms of evaluation research—particularly among the behavioural and social sciences. Such issues are addressed fully in the many specialist publications. However, the basic principles of study design are considered here, along with a small number of issues which are of greatest relevance to the evaluation of health promotion programmes.

Attribution of cause and effect: experimental designs and their problems

A source of many dilemmas and complexities faced by evaluators is the desire to establish a clear relationship between an intervention and a health outcome through a single 'definitive' study. The model in Figure 1 represents an attempt to present this complexity in a visible form. The first level of success is in the achievement of health promotion outcomes, which may be defined as health promotion objectives. Once a reasonable measure of a health promotion objective has been identified (some examples are given above), the next major task is to develop a research design which will allow the use of this measure effectively to determine whether or not an intervention had the intended effects.

There are several essential elements to establishing a relationship between an intervention and an observed outcome. These are thoroughly discussed in existing texts, but can be summarised

as consisting of five essential elements (Windsor *et al.*, 1984):

- a representative sample of the target population or programme recipients;
- one or more pre-tests assessing the objects of intervention;
- an unexposed group for comparison;
- random assignment of the sample to experimental or control groups;
- one or more post-tests to measure effects on the objects of intervention.

Such a design allows for assessment of change by comparing the situation existing before and after intervention. Because individuals have been randomly assigned to intervention and control groups, such a design means that observed change in the study population compared to the control population can be more reliably attributed to the effects of the intervention.

Unfortunately, meeting these basic criteria for the randomised design has proved difficult and often runs counter to the valued processes in health promotion concerning participation in decision-making (Allison and Rootman, 1996). Though some studies have successfully employed this design, most have been narrowly defined, typically restricted to single issues (e.g. smoking), single health promotion objectives (e.g. improving health literacy, changing health behaviour), and interventions undertaken in highly manageable, 'closed' systems such as schools, health clinics, and workplaces. In some cases the volume of studies conforming to these study design criteria has allowed for meta-analysis of results from multiple studies (Mullen *et al.*, 1985, 1992; Kotte *et al.*, 1988; Bruvold, 1993). Such meta-analysis is particularly helpful in improving understanding of this type of intervention, can improve confidence in the validity of findings from individual studies and assess the potential for reproduction.

These studies are important in advancing knowledge and building credibility for health promotion but, for community-based and community-wide programmes, they may be too restrictive, and may ultimately be self-defeating by reducing the effectiveness of the intervention or rendering it impossible to reproduce (Black, 1996). Alternative approaches have to be identified.

Alternatives to experimental design

In circumstances where, for practical reasons (often financial) there are no opportunities to establish a reference population, additional strategies to strengthen inference about programme effects have been developed. These include monitoring changes over time in the object of interest, referred to as a 'time series design'. This is the simplest and least obtrusive form of evaluation. It can often involve use of existing methods of record-keeping; for example, monitoring change in the use of a screening service before, during and after a programme to promote improved uptake; phasing the introduction of interventions into different communities, and observing a change in the intervention population in equivalent phases related to the introduction of the intervention. Such a design temporarily creates a 'non-intervention' population. This is a useful design to overcome the ethical dilemma of deliberately withholding an intervention to a study population. It does not so easily allow for detection of longer-term effects of interventions as a traditional experimental or quasi-experimental design. Differing intervention intensity in different populations is particularly feasible when an intervention consists of different elements (e.g. organisational change, personal education, mass media education). The programme can be offered as a whole to one population, while, by contrast, only the individual component parts are offered to other populations. Green and Lewis (1986) have described a hierarchy of experimental designs, including those above, which provides guidance on the best combinations of the elements of experimental design for varying circumstances.

Strategic issues in evaluating community/ population interventions

Beyond these technical solutions, there is a more fundamental and strategic problem in the use of experimental designs in the evaluation of health promotion programmes. In interventions which are designed to influence human behaviour and social interactions, the artificial assignment of individuals in communities to intervention and control groups is not only often impractical, but frequently impossible as it places quite unrealistic constraints on the intervention design. For example, it is virtually impossible to use the mass media in such a way that the intervention only reaches a randomly selected population group. Further, many health promotion programmes actively draw upon political systems and community networks as part of the intervention. In such circumstances the 'random' allocation of

individuals would place impossible constraints on the possibility of actively using community networks.

As well as these practical constraints, interventions have been strategically designed to influence populations rather than individuals. This 'population' approach to intervention has been impressively articulated by Rose (1985), and advanced scientifically and given profile by several large cardiovascular disease prevention programmes operating in the 1980s (Shea and Basch, 1990). It has become the favoured way of organising comprehensive health promotion programmes to benefit whole populations, through multiple interventions directed towards different health promotion outcomes. In these studies the strategy was directed to achieve mass shifts in risk factor prevalence and change in policy and organisational practice, rather than simply focusing on improving personal health literacy and behaviour modification among defined individuals.

The cardiovascular health promotion programmes provide a good example of efforts to overcome many of the practical problems for evaluation design in programmes directed at whole populations as opposed to individuals. The cardiovascular programmes sought to modify traditional experimental designs in ways which suited the practicalities of the interventions being organised. Whole populations were the 'unit' of intervention, and were matched with equivalent comparison 'units', geographically isolated from the intervention. Thus, the community was the unit of assignment, but the individual remained the unit of observation.

This quasi-experimental design has become the norm for such programmes and has been widely promoted as the best approach to evaluation of community-based programmes. An enhanced version of this quasi-experimental design, the community intervention trial, advocates identification of a large number of separate community 'units' and random allocation of these to intervention and control groups. This evaluation design has been adopted in several well-known studies in the past decade (Jacobs *et al.*, 1986; COMMIT, 1995; Grosskurth *et al.*, 1995) and is considered by some to be the 'only design appropriate for the evaluation of lifestyle interventions that cannot be allocated to individuals' (Murray, 1995).

Community interventions and social movements

Despite this technical progress in developing suitable evaluation designs for well-defined population interventions, the results from the cardiovascular programmes and from the COMMIT smoking cessation trial—the largest experiment with a community intervention trial design—have generally been considered disappointing in terms of their observable net impact on targeted risks. In most cases, positive results have been seen in both intervention and comparison communities. Explanations of these results not only consider the possibility that the interventions may have been insufficiently intense, too brief, or failed to penetrate a sufficient proportion of the population to have had an impact over and above prevailing 'secular trends', but also that the study designs may not have been as useful or sensitive as required for such complex interventions (Mittelmark *et al.*, 1993; Susser, 1995; Fisher, 1995). In addition, some commentators have pointed to poor understanding of the broad research base for interventions (highlighted above), and emphasised the need for 'creative, dedicated, and rigorous social research' to bring about this understanding (Susser, 1995).

One explanation for observed positive results in both intervention and comparison populations is that there has been a high level of 'contamination' between the artificially separated populations. There is good evidence to suggest that this has occurred in some cases (Nutbeam *et al.*, 1993a). But the truth may be more subtle and complex than this. The major changes in smoking behaviour, leisure time physical activity and food choices, which can be observed in both intervention and comparison communities in these and other studies are not 'chance secular trends', but have been achieved through diverse, sustained public health activism over the past three decades. The results of this activism can be observed through simple, regular observational studies, and have been manifest through changing values and behaviour supported by community organisation and, ultimately, law and regulation. These social movements are powerful, and are likely to have overwhelmed the effects of relatively short-term, localised interventions such as those in the cardiovascular and the COMMIT trials.

The WHO-sponsored programmes, such as the Healthy Cities Project and the Health Promoting Schools Project, are more often depicted as social

movements than as tangible 'interventions' of the type described in the cardiovascular programmes (Tsouros, 1995). Social movements take longer to develop, and are less tangible and predictable (and therefore less easily measured and controlled by conventional means) than organised interventions. This is because they draw upon multiple forms of intervention (education, advocacy, facilitation of social mobilisation), often engage the population affected far more directly in decision-making, and rely to a certain extent on opportunism to guide the direction and emphasis of activities. Such an approach to health promotion appears more capable of addressing some of the underlying social and economic determinants of health which require sustained activism, and to offer greater opportunity for community control and empowerment—some of the more important and valued processes and outcomes in health promotion—but is impractical to evaluate using the tightly defined criteria of experimental design (Baum and Brown, 1989).

The dilemma emerging from this analysis is that the more powerful forms of health promotion action are those which appear to be long-term and least easily predicted, controlled and measured by conventional means. Against this, important and valued advances in knowledge and credibility have come from more tightly defined and controlled interventions, which have been evaluated through the application of experimental designs. This tension between the demands for 'scientific rigour' on the one hand, and the advantages in terms of effectiveness and maintenance that come from less-well-defined and community controlled 'movements' has been regularly discussed in the literature (Stevenson and Burke, 1991; Allison and Rootman, 1996).

Advancing knowledge, improving understanding and credibility are extremely important for the relatively new discipline of health promotion. But, an approach to the advancement of knowledge based only on findings from controlled research design also has real dangers—especially when it excludes other forms of evaluation which do not meet experimental design criteria.

Clearly it is nonsense to believe that all other forms of evaluation and experience cannot add to the base of knowledge and understanding in health promotion. The use of experimental designs to assess the success of the less-well-defined forms of social activism indicated above is at best impractical, and more likely is impossible to manage in ways that do not compromise the activity. Other methods have to be used to evaluate the effectiveness of health promotion.

Building evidence using multiple methods and multiple sources of data

Qualitative public health research can provide depth and insight into people's experiences, and the social contexts that strengthen, support or diminish health. This knowledge and insight is important in explaining observed success or failure in any given programme, and essential for the successful replication and dissemination of new ideas.

Despite this, qualitative research is generally undervalued and under used. Part of the reason for this stems from a value system which has evolved among public health researchers (especially those with substantial training in epidemiology and biostatistics) which gives quantitative, experimental research high status, and tends to devalue the importance of research to determine the process of change which may often be qualitative—frequently referred to as 'soft' research. This may be because the methods involved in qualitative research may be less well defined and in many cases simply unfamiliar to researchers used to experimental designs. As a consequence, such methods may either be inappropriately applied or, when properly applied, inappropriately assessed through academic peer review.

Although the methods may be different, qualitative research can be planned and executed with scientific rigour equal to that of quantitative research. Identification of aims, selection and sampling of subjects, method of investigation, and analysis of results can be as well defined and described in qualitative research as in quantitative research (Denzin and Lincoln, 1994).

Rather than imposing impractical and irrelevant evaluation designs, evidence of success in health promotion may best be built on data which are derived from several different sources—some of which may be experimental studies, but many of which will be observational studies, making use of qualitative as well as quantitative information. The search for the 'single definitive study' is illusory and inevitably leads to overly simplistic solutions.

Instead of arguing the relative strengths and weaknesses of quantitative/qualitative research, and experimental/observational research, most researchers involved in the evaluation of health promotion interventions recognise the synergistic

effects of combining different methods to answer different research and evaluation questions (De Vries *et al.*, 1992; Steckler *et al.*, 1992; Baum, 1995).

One promising approach to the use of multiple methods is the concept of research 'triangulation' to improve confidence in research findings. This approach is now well established among qualitative researchers, and involves accumulating evidence from a variety of sources. The logic of this approach is that the more consistent the direction of the evidence produced from different sources, the more reasonable it is to assume that the programme has produced the observed effects. Triangulation simply means using more than one approach to answer the same question. Different types of triangulation can be used (Gifford, 1996), for example:

- **Data source triangulation**, which involves using different kinds of information to investigate a given research question, such as client records, minutes of meetings, published documents, and interviews with key informants.
- **Researcher triangulation**, which involves more than one researcher in data collection and analysis. This approach can be particularly useful if the researchers hold different theoretical and/or methodological perspectives.
- **Methods triangulation**, which involves using a number of different methods, such as focus group discussions, individual interviews, observation of meetings and other interactions, to investigate a nominated issue.

The use of 'triangulation' has much merit in the evaluation of health promotion, especially where experimental research design may be inappropriate, impractical, or provide only part of the picture in a multi-level intervention. Combining information from different quantitative and qualitative sources to assess for consistency in results can provide powerful evidence of success, as well as providing insight to the processes of change in populations and organisations.

CREATING CONDITIONS FOR SUCCESS: EVALUATION OF PROCESS

By recognising the benefits of combining different research methods to answer different research questions, the distinction between 'outcome' and 'process' evaluation indicated in Figure 2 becomes somewhat blurred. Understanding the process of an intervention (or social movement) is of great importance in its own right, but is also essential to build the evidence on which 'success' is determined. Investigation of how a programme is implemented, what activities occurred under what conditions, by whom, and with what level of effort, will ensure that much more is learned and understood about success or failure in achieving defined outcomes. Through this understanding it is possible to identify the conditions which need to be created to achieve successful outcomes.

A number of basic, and inter-related process evaluation aims can be identified in published work. These are considered below.

Programme reach: did the programme reach all of the target population?

In any health programme, a key element of success has to be in achieving optimal contact with the defined target population—whether this is an 'at-risk' group, a whole community, managers in an organisation, or community leaders/politicians. To evaluate the effects of a programme, it is essential to be able to determine the extent and level of exposure to it.

This is relatively simple where the intervention can be clearly defined—for example attending a smoking cessation group, receiving a pamphlet or media communication (Cumming *et al.*, 1989)—but far more difficult in community programmes where the intervention is less easy to define, and determining exposure a far more complex task. Methods which have been used to measure programme exposure range from simple audit and record keeping, to sophisticated monitoring among defined groups.

The heart health programmes in the US referred to above all developed sophisticated systems for monitoring population exposure (Flora *et al.*, 1993). In these programmes, exposure was monitored through a variety of methods, including the use of specially designed contact cards which were completed by everyone who participated in the intervention. The data were used to determine the demographic profiles of participants, document each participant's total number of exposures to the intervention, refine and target intervention programmes, assess the immediate and long-term impacts of the interventions through follow-up surveys, and provide a historical record of the entire intervention effort.

Other studies of programme reach have

explored awareness of interventions among target populations in communities, programme gatekeepers (for example teachers, general practitioners) (Perhats *et al.*, 1996) and within organisations such as schools, and in worksites (Fielding, 1990; Smith *et al.*, 1993).

Programme acceptability: is the programme acceptable to the target population?

Although a programme may reach its intended audience, the response of this population to the programme is critical. Studies which assess the acceptability of programmes and their subsequent 'ownership' by the target population, and/or the programme 'gatekeeper' (teacher, health worker, manager, politician) form an essential part of process evaluation (Bracht *et al.*, 1994). There are different dimensions to this question which examine the extent to which people feel involved in a programme, able to influence its direction and outcome; perceptions of the relevance of the programme to people's needs and priorities; and perceptions of the feasibility of actions which are advocated through a programme.

Studies of 'gatekeepers' have looked at their experiences of implementing programmes, the acceptability of different programme activities, the perceived effects of projects, and suggestions for modifications. Examples of such studies can be found with professional groups, particularly teachers and doctors (Newman and Nutbeam, 1989; Arbeit *et al.*, 1991; Murphy and Smith, 1993). Less common in published reports are evaluations which have taken the views and experiences of communities into account.

Programme integrity: was the programme implemented as planned?

Finally, in order to fully understand observed change in health promotion outcomes, it is essential to record the extent to which a programme was implemented as planned. Failure to achieve defined programme objectives could be a result of a poor intervention, or a poorly executed intervention. Interventions which have been evaluated and determined as effective by a group of highly motivated researchers working with equally motivated volunteers, are not automatically well received, executed and sustained when translated into 'real-life' settings (Nutbeam *et al.*, 1993b).

Observing and recording activities is the simplest method of doing this. More sophisticated forms of analysis of programme integrity may involve tracing the 'chain of events' within a discreet community, determining such issues as dilution or distortion of programme inputs (Scheirer *et al.*, 1995). This approach has been used in community-based programmes to understand the dynamics of intervention implementation within defined social or professional networks, or in specific settings such as schools (Sobol *et al.*, 1989; Perry *et al.*, 1990; Russos *et al.*, 1997).

One of the most sophisticated examples of comprehensive process evaluation was that employed by the Child and Adolescent Trial for Cardiovascular Health (CATCH) (McGraw *et al.*, 1994). The purpose of this effort was to 'describe the implementation of the programme, quality control and monitoring, and explain programme effects' (Stone, 1994). These aims summarise the purpose of good process evaluation as a complementary task to evaluation research primarily directed towards measuring outcomes. At one level, process evaluation can support and enhance causal inference in studies. At another level, it opens the door through which basic experimental studies can be repeated, refined and widely disseminated by defining the conditions which need to be created for success in achieving programme objectives. In this way, process evaluation has particular relevance to policy-makers and practitioners.

CONCLUDING REMARKS: EVALUATION IN A COMPLEX ENVIRONMENT

Evaluation of health promotion is a difficult enterprise which is often done poorly. Many of the problems faced by practitioners attempting to evaluate health promotion activity stem from unreasonable expectations of both the activity and the evaluation. Health promotion is a complex field. Tracing the causal path from a community intervention to subsequent long-term changes in health and social outcomes is fraught with difficulty, and it is inappropriate and unrealistic in most cases for programmes to be expected to do this. Far more relevant is for health promotion interventions to be judged on their ability to achieve the health promotion outcomes defined above, using evaluation methods which best fit the activity. Such a position does not always fit comfortably with prevailing views in the health and medical

community on what constitutes 'rigorous' method and 'hard' evidence.

Four key challenges emerge from this analysis:

(i) *Using research evidence more systematically in the planning of activities.* The volume of research to support health promotion action has grown remarkably in the past 20 years and needs to be applied to current health promotion practice. In particular, this research evidence should be used more creatively to improve understanding of the complexity of relationships between the different levels of outcome illustrated in Figure 1, and to provide greater insight into the definition of problems to be addressed and the interventions required to address them. A wide range of research evidence needs to be systematically incorporated into activity planning.

(ii) *Improving the definition and measurement of outcome*. Poor definition of programme objectives—whether these are expressed in terms of valued outcomes and/or valued processes—often leads to inappropriate expectations concerning evaluation and accountability. For the future, it is essential that programme objectives are more clearly defined, and that relevant and sensitive measures are used to assess progress in achieving these objectives. This will require more systematic development and use of valid and reliable indicators of health promotion outcomes, particularly measures of social mobilisation, public policy and organisational practice.

(iii) *Adopting appropriate evaluation intensity*. Not all programmes need to be evaluated to the same level of intensity or using the same evaluation designs. The hierarchy in Figure 2 indicates how the evaluation question changes with the evolution of a programme. It suggests that those programmes which are truly innovative, testing for the first time potentially costly, controversial, or otherwise risky forms of intervention, need close scrutiny and the most structured and comprehensive approaches to evaluation. Those which have previously been shown to work in a variety of circumstances, that are low cost and low risk, will require more modest monitoring for the purposes of accountability and quality control.

(iv) *Adopting appropriate evaluation design*. There has been an unrealistic expectation to adopt experimental research designs which have been developed for medical research. This is inappropriate at several levels. Firstly, the constraints on the intervention strategy imposed by such experimental designs make it virtually impossible to use the community-based approaches which are considered to be the most valuable and effective. Secondly, experimental designs have been shown to be deficient as an evaluation tool for complex and multi-dimensional activities. Thirdly, because they are such powerful and persuasive scientific tools, randomised controlled trials for outcome evaluation have tended to eclipse the value and relevance of other research methods for outcome evaluation—especially qualitative methods—and for evaluating the process of change.

For the future, it is important to foster and develop feasible evaluation designs which combine different research methodologies, quantitative with qualitative. The generation and use of a diverse range of data and information sources will generally provide more illuminating, relevant and sensitive evidence of effects than a single 'definitive' study. Process evaluation not only provides valuable information on how a programme is implemented, what activities occur under what conditions, by whom, and with what level of effort, but will also ensure that much more is learned and understood about success or failure in achieving defined outcomes. Through this understanding it is possible to identify the conditions which need to be created to achieve successful outcomes. Evaluations have to be tailored to suit the activity and circumstances of individual programmes—no single method can be 'right' for all programmes.

Address for correspondence:
Don Nutbeam
Department of Public Health
University of Sydney
Sydney
NSW 2006
Australia

REFERENCES

Abelin, T., Brzezinski, Z. J. and Carstairs, V. D. L. (1987) *Measurement in Health Promotion and Protection*, WHO

Regional Publications, European Series No. 22. WHO, Copenhagen.

Allensworth, D. D. (1994) The research base for innovative practices in school health education at the secondary level. *Journal of School Health*, **64,** 180–187.

Allison, K. R. and Rootman, I. (1996) Scientific rigour and community participation in health promotion research: are they compatible? *Health Promotion International*, **11,** 333–340.

Arbeit, M. L., Serpas, D. C. and Johnson, C. C. (1991) The implementation of a cardiovascular school health promotion program: utilisation and impact of a school health advisory committee: The Heartstart program. *Health Education Research*, **6,** 423–440.

Badura. B. and Kickbusch, I. (1991) *Health Promotion Research: Towards a New Social Epidemiology*, WHO Regional Publications, European Series No. 37. WHO, Copenhagen.

Baum, F. (1995) Researching public health: beyond the qualitative–quantitative method debate. *Social Science and Medicine*, **55,** 459–468.

Baum, F. and Brown, V. (1989) Healthy Cities (Australia) Project: issues of evaluation for the new public health. *Community Health Studies*, **8,** 140–149.

Black, N. (1996) Why we need observational studies to evaluate the effectiveness of health care. *British Medical Journal*, **312,** 1215–1218.

Bracht, N., Finnegan, J. R., Rissel, C., *et al.* (1994) Community ownership and program continuation following a health demonstration project. *Health Education Research*, **9,** 243–255.

Bruvold, W. H. (1993) A meta-analysis of adolescent smoking prevention programs. *American Journal of Public Health*, **86,** 872–880.

Catford, J. C. (1993) Auditing health promotion: what are the vital signs of quality? (Editorial). *Health Promotion International*, **8,** 67–68.

Chapman, S. (1993) Unravelling gossamer with boxing gloves: problems in explaining the decline in smoking. *British Medical Journal*, **307,** 429–432.

Cheadle, A. (1992) Environmental indicators: a tool for evaluating community-based health promotion programs. *American Journal of Preventive Medicine*, **8,** 345–350.

Cohen, D. (1994) Health promotion and cost-effectiveness. *Health Promotion International*, **9,** 281–287.

COMMIT Research Group (1995) Community Intervention Trial for Smoking Cessation (COMMIT): I. Cohort results from a four year community intervention. *American Journal of Public Health*, **85,** 183–192.

Coppel, S., King, L., Stoker, L., *et al.* (1994) *Program Management Guidelines for Health Promotion.* NSW Health Department, Sydney.

Cumming, M. R., Sciandra, R., Davis, S. and Rimer, B. (1989) Response to anti-smoking campaign aimed at mothers with young children. *Health Education Research*, **4,** 429–437.

De Vellis, R. F., Alfieri, W. S. and Ahluwalia, G. (1995) The importance of careful measurement in health education research, theory and practice (Editorial). *Health Education Research*, **10**, i–iv. [This edition of the journal (Vol. 10, No. 1) was a special issue on measurement in health education.]

De Vries, H., Weijts, W., Dijkstra, M. and Kok, G. (1992) The utilisation of qualitative and quantitative data for health education program planning, implementation and evaluation: a spiral approach. *Health Education Quarterly*, **19,** 101–115.

Denzin, N. K. and Lincoln, Y. (eds) (1994) *Handbook of Qualitative Research*. Sage, Thousand Oaks, CA.

Epstein, A. M. (1990) The outcomes movement—will it get us where we want to go? *New England Journal of Medicine*, **323,** 266–270.

Evans, R. G., Barer, M. L. and Marmor, T. R. (eds) (1994) *Why Are Some People Healthy and Others Not? The Determinants of Health in Populations*. Walter de Gruyter, New York.

Fielding, J. E. (1990) Worksite health promotion programs in the United States: progress, lessons and challenges. *Health Promotion International*, **5,** 75–84.

Fisher, E. B (1995) The results of the COMMIT Trial. *American Journal of Public Health*, **85,** 159–160.

Flay, B. (1986) Efficacy and effectiveness trials (and other phases of research) in the development of health education programs. *Preventive Medicine*, **15,** 451–474.

Flora, J. A., Lefebre, R. C., Murray, D. M., *et al.* (1993) A community education monitoring system: methods from the Stanford Five-Cities Program, the Minnesota Heart Health Program, and the Pawtuckett Heart Health Program. *Health Education Research*, **8,** 81–95.

Gifford, S. (1996) Qualitative research: the soft option? *Health Promotion Journal of Australia*, **6,** 58–61.

Goodman, R. M. and Steckler, A. B. (1987) A model for institutionalisation of health promotion programs. *Family and Community Health*, **11,** 63–78.

Goodman, R., McLeroy, K. R., Steckler, A. B. and Hoyle, R. H. (1993) Development of level of institutionalisation scales for health promotion programs. *Health Education Quarterly*, **20,** 161–178.

Green, L. W. (1987) Three ways research influences practice: the public's right to know and the scientist's responsibility to educate. *Health Education*, **4,** 44–49.

Green, L. W. and Lewis, F. M. (1986) *Measurement and Evaluation in Health Education and Health Promotion*. Mayfield, Palo Alto, CA.

Grosskurth, H., Mosha, F., Todd, J., *et al.* (1995) Impact of improved STDs on HIV infection in rural Tanzania. *Lancet*, **346,** 530–536.

Hawe, P., Degeling, D. and Hall, J. (1990) *Evaluating Health Promotion: A Health Workers Guide*. Maclennan & Petty, Sydney.

Hawe, P., Noort, M., King, L. and Jordens, C. (1997) Multiplying health gains: the critical role of capacity building within health promotion programs. *Health Policy*, **39,** 29–42.

Holman, C. D. J., Donovan, R. J. and Corti, B. (1993) Evaluating projects funded by the West Australian Health Promotion Foundation: a systematic approach. *Health Promotion International*, **8,** 199–208.

Jackson, C., Fortmann, S. P., Flora, J. A., *et al.* (1994) The capacity-building approach to intervention maintenance implemented by the Stanford Heart Disease Prevention Project. *Health Education Research*, **9,** 385–396.

Jacobs, D. R., Leupker, R. V., Mittelmark, M., *et al.* (1986) Community-wide prevention strategies: evaluation design of the Minnesota Heart Health Program. *Journal of Chronic Disease*, **39,** 775–788.

Kaplan, G. A., Pamuk, E. R., Lynch, J. W., *et al.* (1996) Inequality in income and mortality in the US: analysis of mortality and potential pathways. *British Medical Journal*, **312,** 999–1003.

Kar, S. B., Colman, W., Bertolli, J. and Berkanovic, E. (1988) Indicators of individual and community action for health promotion. *Health Promotion*, **3,** 59–66.

King, L. (1996) An outcomes hierarchy for health promotion: a tool for policy, planning an evaluation. *Health Promotion Journal of Australia*, **6,** 50–51.

Kotte, T., Battista, R., DeFreise, G. and Brekke, M. (1988) Attributes of successful smoking interventions in medical practice: a meta analysis of 39 controlled trials. *Journal of the American Medical Association*, **259,** 2888–2889.

Macdonald, G., Veen, C. and Tones, K. (1996) Evidence for success in health promotion: suggestions for improvement. *Health Education Research*, **11,** 367–376.

McGraw, S., Stone, E., Osganian, S. K., *et al.* (1994) Design of process evaluation within the Child and Adolescent Trial for Cardiovascular health (CATCH). *Health Education Quarterly*, **S2,** S5–S26.

Mittelmark, M. B., Hunt, M. K., Heath, G. W. and Schmid, T. L. (1993) Realistic outcomes: lessons from community-based research and demonstration programs for the prevention of cardiovascular diseases. *Journal of Public Health Policy*, **14,** 455–462.

Mullen, P. D., Green, L. W. and Persinger, G. (1985) Clinical trials of patient education for chronic conditions: a comparative meta-analysis. *Preventive Medicine*, **14,** 753–781.

Mullen, P. D., Mains, D. A. and Velez, R. (1992) A meta-analysis of controlled trials in cardiac patient education. *Patient Education and Counselling*, **14,** 143–162.

Murphy, S. and Smith, C. (1993) Crutches, confetti or useful tools: professional views on and uses of health education leaflets. *Health Education Research*, **8,** 205–215.

Murray, D. M. (1995) Design and analysis of community trials: lessons from the Minnesota Heart Health Program. *American Journal of Epidemiology*, **142,** 569–575.

Newman, R. and Nutbeam, D. (1989) Teachers views of the Family Smoking Education Project. *Health Education Journal*, **48,** 9–13.

Noack, H. (1988) Measuring health behaviour and health: towards new health promotion indicators (Editorial). *Health Promotion*, **3,** 5–12. [This edition of the journal (Vol. 3, No. 1) was devoted to health promotion indicators.]

Nutbeam, D. (1996a) Achieving 'best practice' in health promotion: improving the fit between research and practice. *Health Education Research*, **11,** 317–326.

Nutbeam, D. (1996b) Health outcomes and health promotion: defining success in health promotion. *Health Promotion Journal of Australia*, **6,** 58–60.

Nutbeam, D., Smith, C. and Catford, J. (1990) Evaluation in health education, progress, problems and possibilities. *Journal of Epidemiology and Community Health*, **44**, 83–89.

Nutbeam, D., Smith, C., Murphy, S. and Catford, J. (1993a) Maintaining evaluation designs in long-term community based health promotion programs. *Journal of Epidemiology and Community Health*, **47,** 123–127.

Nutbeam, D., Macaskill, P., Smith, C., *et al.* (1993b) Evaluation of two school smoking education programs under normal classroom conditions. *British Medical Journal*, **306,** 102–107.

Oldenburg, B., Hardcastle, D. and Ffrench, M. (1996) How does research contribute to evidence-based practice in health promotion? *Health Promotion Journal of Australia*, **6,** 15–20.

Orlandi, M. (1986) The diffusion and adoption of worksite health promotion innovations: an analysis of barriers. *Preventive Medicine*, **15**, 522–536.

Paine, A. L. and Fawcett, S. B. (1993) A methodology for monitoring and evaluating community health coalitions. *Health Education Research*, **8,** 403–416.

Parcel, G. S., Taylor, W. C., Brink, S. G., *et al.* (1989) Translating theory into practice: intervention strategies for the diffusion of a health promotion innovation. *Family and Community Health*, **12,** 1–13.

Perhats, C., Oh, K., Levy, S., *et al.* (1996) Role differences and gatekeeper perceptions of school-based drug and sexuality education programs. *Health Education Research*, **11,** 11–27.

Perry, C. L., Murray, D. M. and Griffen, G. (1990) Evaluating statewide dissemination of smoking prevention curricula: factors in teacher compliance. *Journal of School Health*, **60,** 501–504.

Rissel, C., Finnegan, J. and Bracht, N. (1995) Evaluating quality and sustainability: issues and insights from the Minnesota Heart Health Program. *Health Promotion International*, **10,** 199–207.

Rose, G. (1985) Sick individuals and sick populations. *International Journal of Epidemiology*, **14,** 32–38.

Russos, S., Hovell, M. and Keating, K. (1997) Clinician compliance with primary prevention of tobacco use: the impact of social contingencies. *Preventive Medicine*, **26,** 44–52.

Rychetnik, L., Nutbeam, D. and Hawe, P. (1997) Lessons from a review of publications in health promotion journals 1989–94. *Health Education Research*, in press.

Sanderson, C., Haglund, B., Tillgren, P., *et al.* (1996) Effect and stage models in community intervention programs: and the development of the Model for Management of Intervention Program Preparation. *Health Promotion International*, **11,** 143–154.

Sanson-Fisher, R., Redman, S., Hancock, L., *et al.* (1996) Developing methodologies for evaluating community-wide health promotion. *Health Promotion International*, **11,** 227–236.

Scheirer, M. A., Shediac, M. C. and Cassady, C. E. (1995) Measuring the implementation of health promotion programs: the case of the Breast and Cervical Cancer Program in Maryland. *Health Education Research*, **10,** 11–25.

Schultz, A. J., Israel, B., Zimmerman, M. A. and Checkoway, B. N. (1995) Empowerment as a multi-level construct: perceived control at the individual, organizational and community levels. *Health Education Research*, **10,** 309–327.

Schwarz, R., Goodman, R. and Steckler, A. (1995) Policy advocacy interventions for health promotion and education: advancing the state of practice (Editorial). *Health Education Quarterly*, **22,** 421–426.

Shea, S. and Basch, C. E. (1990) A review of five major community-based cardiovascular disease prevention programs. Part II: intervention strategies, evaluation methods and results. *American Journal of Health Promotion*, **4,** 279–287.

Sheill, A. (1997) Health outcomes are about choices and values: an economic perspective on the health outcomes movement. *Health Policy*, **39,** 5–15.

Smith, C., Nutbeam, D., Roberts, C. and Macdonald, G. (1993) The health promoting school: progress and future challenges in Welsh secondary schools. *Health Promotion International*, **7,** 171–179.

Sobol, D. F., Rorhbach, L. A., Dent, C. W., *et al.* (1989) The integrity of smoking prevention curriculum delivery. *Health Education Research*, **4,** 59–67.

Steckler, A., McLeroy, K. R. and Goodman, R. M. (1992) Towards integrating qualitative and quantitative methods:

an introduction (Editorial). *Health Education Quarterly*, **19,** 1–8. [This edition of the journal (Vol. 19, No. 1) was devoted to examining issues concerning the use of qualitative and quantitative research in health education.]

Stevenson, M. and Burke, M. (1991) Bureaucratic logic in new social movement clothing: the limits of health promotion research. *Health Promotion International*, **6,** 281–289.

Stone, E. (1994) Foreword to Special Supplement on the process evaluation in the multicenter Child and Adolescent Trial for Cardiovascular Health (CATCH). *Health Education Quarterly*, **S2,** S3–S4.

Susser, M. (1995) The tribulations of trials: intervention in communities. *American Journal of Public Health*, **85,** 158.

Syme, S. L. (1996) Rethinking disease: where do we go from here? *Annals of Epidemiology*, **6,** 463–468.

Tones, K. (1992) Measuring success in health promotion: selecting indicators of performance. *Hygie*, **11,** 10–14.

Tsouros, A. G. (1995) The WHO Healthy Cities Project: state of the art and future plans. *Health Promotion International*, **10,** 133–141.

Van Driel, W. G. and Keijsers, J. F. (1997) An instrument for reviewing the effectiveness of health education and health promotion. *Patient Education and Counselling*, **30,** 7–17.

Viney, R. (1996) Health promotion, economics, priority setting, resource allocation and outcomes at the margin. *Health Promotion Journal of Australia*, **6,** 9–14.

Weir, R. D. (1991) Research in public health: who says; who does; and who cares. *Journal of Public Health Medicine*, **13,** 151–157.

Windsor, R. A., Baranowski, T., Clark, N. and Cutter, G. (1984) *Evaluation of Health Promotion and Health Education Programs*. Mayfield, Palo Alto, CA.

World Health Organization (1986) *Ottawa Charter for Health Promotion*. World Health Organization, Geneva.