

programme staff's sense of 'certainty' referred to above and anecdotal evidence of effectiveness are, on their own, certainly not going to be enough.

Also common is a converse 'impression', or unsupported belief, that nothing is working and that a programme's interventions are futile within the wider problems of society. This perception is equally unreliable – and undesirable in terms of staff motivation. Systematic evaluation can be just as helpful in showing that certain outcomes are being achieved, and in countering a sense of futility in staff.

Thus, as opposed to a process of forming chance or subjective impressions, programme evaluation is ultimately a conscious and systematic process. This is evident in the following definition: 'Program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming' (Patton, 1997, p. 23). In drawing the distinction between unsystematic judgments and opinions about programme performance and programme evaluation, Newcomer, Hatry and Wholey (1994, p. 3) say: 'In the public arena and in the private nonprofit sector, there is always a cacophony of unsystematic feedback on program performance. *Program evaluation* is the systematic assessment of program results and, to the extent feasible, systematic assessment of the extent to which the program caused those results.'

THE LOGIC OF PROGRAMMES

There is a certain logic to programmes and to programme evaluation, awareness of which may be of assistance to programme managers and staff. Making this logic explicit, and examining the sequence of events of which it is comprised, may serve to clarify many difficulties that are commonly experienced with programme evaluation. Put in the simplest terms possible, the logic consists of the sequence of components represented in Figure 1.

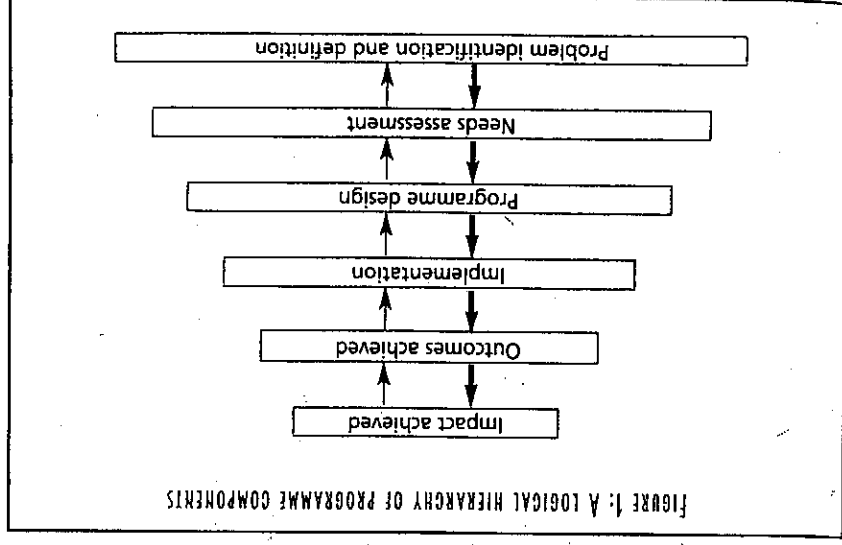


FIGURE 1: A LOGICAL HIERARCHY OF PROGRAMME COMPONENTS

IMPROVING PRACTICE THROUGH EVALUATION

JOHANN LOUW

The notion of evaluation is often not appealing to programmes, their management staff and service deliverers. In the face of sometimes overwhelming need, programme staff want to get on with alleviating that need. They do not want to spend time and money on an evaluation exercise which may place the programme at risk, or which may reflect negatively on programme staff. Yet the practice of evaluating one's own efforts is part of everyday life, of trying to do things better, and of improving the performance of even the most routine aspects of our activities.

How to conduct an evaluation? Rather, it is an attempt to answer the question 'How can evaluation be used to strengthen or improve interventions?' In the light of the tremendous demands made on human services in South Africa, especially in relation to children, evaluation will not serve a useful function if it is seen as something threatening and punitive, or as a trial to be endured and then forgotten. However, programme staff are generally interested in improving their performance and would like to establish whether their activities are making a difference. If they are convinced that evaluation can help in this, they may begin to incorporate an evaluative way of thinking into their everyday activities. Even better, they can be encouraged to integrate evaluation into programme processes, to strengthen the intervention and to make the evaluation a built-in part of it.

Programme staff are themselves frequently certain that their interventions are effective. They know how relevant their work is, and they see evidence of this in feedback from the people with whom they work. Despite this sense of making a difference, programme activities and results have to be accounted for to a wider audience. Staff need to be able to provide documented evidence supporting their claim that the programme is effective. Obviously, this evidence must be convincing, otherwise funding agencies, policy makers, and the general public will simply not believe it. The pro-

ful in informing programme decisions - by any stakeholders, from clients to funders - is lack of adequate consideration of what happens (or happened) at the beginning of the process. Taking the field of child health and nutrition as an example, Chapter 2 identifies a set of immediate causes of malnutrition. These include inadequate dietary intake, disease and psychosocial stress and trauma. In addition, underlying determinants of malnutrition are identified as food security, access to health services, care of women and children, and education and information.

In relation to such an analysis of causes, two considerations are important - to understand, theoretically, the practical problem that is faced, and how the intervention might act as a change mechanism. Evaluators will ask programme planners and managers to think about the question 'Is the programme intent appropriate to the problem it wishes to address?' This requires an understanding of the nature of the social problem that is being addressed and the processes that produce it (Chen & Ross, 1989). Social science knowledge provides numerous pointers to the social causes of problems. Evaluators familiar with this literature (e.g. on the psychological effects of malnutrition) can provide important support to planners in understanding the problem, and in designing a programme which has a reasonable chance of success. This is sometimes referred to as problem theory. The problem is defined in as exact terms as possible, possible causes are specified, and the persons or groups who are typically affected are identified. This is not really an area for programme evaluators as such, but more for specialists in the relevant content areas. Nevertheless, evaluators can play an important part in identifying problems and describing them empirically. They can also help by directing attention to the relevant literature on the nature and causes of social problems.

When evaluators talk of research and theory, it is not unusual for programme staff to say that they are not in the theory business or that they are practical people interested in making a difference, and not interested in theoretical or research issues. But this is a false dichotomy. Take for example the importance of knowing the literature on a particular problem area. Literature searches and reviews are playing an increasingly important role in the accumulation of knowledge, not only in the social sciences, but also in the world of practice. The importance for social scientists as well as practitioners of knowing the literature in a specific problem area cannot be over-emphasised. For a start, it makes it less likely that what others have tried already will be duplicated. This is the well-known 'reinventing the wheel' accusation. In addition, it integrates and summarises the state of knowledge in a given domain, allowing overall conclusions to be drawn from many separate studies. Finally, knowing the literature often allows implications for policy and practice - such as establishing the conditions under which an intervention is likely to be more or less successful - to be clearly stated and understood.

'Research' at the early stages of programme development need not include only literature searches. It can also be as practical and down-to-earth as searching for programmes and initiatives in the same problem domain, with objectives similar to one's own. This draws attention to the importance of networking and of sharing information between different agencies and service deliverers, so that all can benefit from learning through the experience of others.

Needs assessment

In more specific terms, part of the process of problem identification is to ask questions about the assessment of need, to assist in the selection of problems on which to focus resources. Needs can be defined as those things without which, or without a sufficient amount of which, a negative or undesirable state occurs (Lipsey, Wilson, Shayne, Dertson & Newbrough, 1996). A negative state may be defined objectively, such as poor health, or subjectively through social construct. Negative states of a social variety are the sorts of problem commonly addressed by social services. Part of a needs assessment therefore is to identify and rank the gaps between problem areas and the social services available to address them ('gaps analysis'). For Lipsey *et al.* (1996) this definition helps to avoid the confusion of solutions with needs. We do not need psychotherapy; we need psychological well-being. It is the absence of the latter which is an undesirable state. Psychotherapy may well be a means of achieving psychological well-being, but it is not the need.

If the evaluator takes part in the planning of a new programme (and it is clear that evaluations are most useful if they are built into the planning process), he or she will ask about the unmet needs to be addressed by the programme. Needs assessment is a sub-discipline of evaluation in its own right, and has developed a methodology of its own. It relies on many sources of information to describe as accurately and reliably as possible the 'things without which a negative or undesirable state occurs'. The sources may include social indicators, community surveys, site visits and situation analyses, key informant interviews, and rapid appraisal techniques - several of which are illustrated in the chapters which follow.

Needs assessment, done before the programme starts, also provides a baseline measure against which the situation, after completion of the programme, can be compared. If parasitic infection among 4-6-year-old children in a particular region was defined as a problem, and the need to address it was identified, it would help to be able to say that prior to the programme 80% of the children were affected, but after the programme only 60% of the children were still infected.

Needs assessments are also important in planning the delivery of services. If the assessment has been thorough, it enhances the chances of reaching and serving the population of interest. Thus the evaluation described in Chapter 7 identified the lack of a 'systematic and rigorous needs assessment survey before embarking on an actual implementation' as one of the major oversights in the planning of the intervention.

Assessing the programme design

Generally, more attention needs to be paid to the question 'Why should this intervention work?' With new programmes in particular, special attention ought to be paid to understanding how the intervention is supposed to bring about the desired changes. In Chapter 4 for instance, the authors base their intervention of a school-feeding scheme on 'sound conceptual and empirical evidence linking nutrition and school performance'. In this case, therefore, there is a clear theoretical basis for expecting the intervention to have its desired effects.

Monitoring implementation

If we return to the example of a school-feeding scheme provided earlier, we can see why a focus on implementing the programme is important. Many things can go wrong with the treatment plan of delivering food parcels to children under the age of 7. People may defraud the scheme of either its funding or the actual food, children or parents may not accept the type of food provided, and the schools, as delivery points, may not have the administrative capacity to carry the intervention through effectively. But anticipating what might happen enables us to track the contextual factors operating on the programme, and may eventually lead us to take other steps in order for the children to benefit.

Just as programme personnel may have espoused theories of action and theories-in-use, so we can expect discrepancies between the programme-as-designed and the programme-as-delivered. Not only that, but there may very well be variations in delivery between the different sites at which the programme is delivered. Thus the programme cannot be regarded as a single entity delivered in a uniform way in all settings. Posavac and Carey (1997, p. 121) describe monitoring the operation of programmes as 'the most fundamental form of program evaluation', and refer to it as 'an examination of the program itself - its activities, the population it serves, how it functions, and the condition of its participants'. Indeed, de Wilde's (1967) collected studies showed that programme implementation and administration were of critical importance in developing countries. If service delivery in developing countries is very poor, as these studies suggest, it makes it impossible to say anything about impact, since poor delivery or implementation virtually guarantees that the programme will fail. It would be unfair in such a situation to conclude that the intervention is ineffective, since it has not been given a proper opportunity to bring about the desired change. This lack of valid evaluation data is one reason why it is difficult to design more effective projects in developing countries.

Why should this be the case? Several reasons can be given for the importance of monitoring implementation or of 'process evaluation' (as it is sometimes called) in the overall logic of the programme and its evaluation. In the most extreme case, this part of the evaluation enterprise provides evidence that planned activities have actually taken place. This information is important for purposes of accountability to funders and others. In less extreme terms, it is important to know not only that the programme has been implemented, but that it has been implemented at a strength that will make it likely to have an effect. Weak implementation is often identified as a major cause of failure of the programme to bring about change. A potentially strong treatment may be built into the design phase, but if it is weakly implemented, it has little chance of being effective. As indicated already, we also need to know whether the programme was delivered uniformly to the appropriate recipients at the right time, or whether delivery was degraded in some way. Xeaton and Sechrest (1981) have referred to this as the 'integrity of treatment'. The best safeguard to ensure that strength and integrity of service delivery are adequately monitored is to actually measure or determine the level of treatment received by participants. We need to have a description of what happened in the delivery of the actual intervention and its activities. For example, in the

However, practitioners are not always aware of their programme theory, and evaluators may serve to help them articulate it. Argyris (1982, p. 83) introduced the notion of a theory of action: 'We begin with the proposition that people hold theories of action about how to produce consequences they intend. Such theories are theories about human effectiveness. By *effectiveness* we mean the degree to which people produce their intended consequences in ways that make it likely that they will continue to produce intended consequences.' Argyris and Schon (1978) earlier distinguished between espoused theories of action (what people say or believe is the theory underlying their actions) and theories-in-use (the bases on which they act). Patton (1997) suggests that the evaluator assist practitioners to articulate their espoused theories of action while, at the same time, continuing to test these against their actual theories-in-use, so that differences may be identified and dealt with. According to Lipsey, treatment or intervention theory is 'a set of propositions regarding what goes on in the black box during the transformation of input to output, that is, how a bad situation is transformed into a better one through treatment inputs' (1993, p. 11). Such a treatment theory would specify at least the following elements: a problem definition that articulates what condition is treatable, for which populations, and under what conditions; the necessary inputs and the interrelationship between them; the important links in the transformation process, and the mediating variables in the process; the expected outputs; and other elements, such as contextual or environmental factors which may affect the treatment process. The important point here is that programme design is more effective if assumptions about the best way to bring about changes (and assumptions about the causes of the problems) are made explicit. Many educational programmes, for example, assume a particular model of change - that if you give people knowledge about the condition and about what they have to do, their attitudes will change, and the attitude change will lead to a change in behaviour. Thus, if faced with the problem of malnutrition, an educational campaign may be launched to inform mothers in poverty-stricken areas on how to prepare nutritious meals from very basic ingredients. In this case the implicit treatment theory may be that the increase in nutritional knowledge brought about through the educational campaign will lead to a change in attitude regarding 'basic ingredients'. In turn, the theory might hold that the change in attitude would lead to the mothers cooking differently and more nutritiously. Such a theory of change could be challenged along a number of dimensions. For a start, it is very hard to change human behaviour by information input alone. Media campaigns to inform people about practising safer sex to reduce the risk of exposure to HIV/AIDS may very well increase knowledge and awareness (see Coyle, Boruch & Turner, 1991), but whether such campaigns lead to change in sexual behaviour is very much in question. Certainly, theorists like Fishbein and Ajzen (1975) and Bandura (1977) have demonstrated quite forcefully that the knowledge-attitude-behaviour sequence of supposed change does not necessarily hold true. Despite this, programmes are frequently designed in which it is assumed that information or knowledge will lead to a change in people's behaviour. Thus, not being clear about one's programme theory, or following an implausible theory, has the very practical consequence of decreasing the chances of the programme making a difference to the condition of its recipients.

ultimate effects. Let us call these 'impacts', also after Linney and Wandersman, these changes observable only at the community level, such as changes in social indicators (like the infant mortality rate and the human development index quoted in Chapter 2). These are clearly effects to be achieved only after a number of intermediate outcomes have been achieved, such as reducing parasitic infection, increasing caloric intake, and increasing school attendance. If the outcomes have been achieved, one may expect the infant mortality rates to decrease and the human development index to rise. Why is this distinction important, and how can it help programme staff? For a start, the distinction clarifies our expectations of what should happen as a result of the intervention and when. It might also guide us as to what to try to measure, and at what phase of the whole programme. Returning to the example of a school-feeding scheme, if the implementation has been closely monitored and has been delivered at full strength, we may expect improvements in the children's health (measured, for example, as an increase in body weight) fairly soon. The improved nutrition should then increase their ability to pay attention in school (which could be measured through systematic observation of their classroom behaviour). Finally, this should lead to more effective learning (which could be gauged through general measures of school achievement). Chapter 4 contains a description of such a causal chain.

Another example could be found in a job skills training programme delivered to groups of unemployed youth. We may ask whether the intended skills were actually acquired as an immediate outcome. But the question of whether those skills enabled the young people to find employment is an impact question. This would have to be asked further down the line. If the outcomes (job skills) have not been achieved, however, there is less chance that the impacts (employment) will be achieved (assuming, of course, that the treatment-theory is valid!). The outcome question is therefore an important one to ask.

Furthermore, clarifying the distinction helps to protect the programme against unrealistic expectations. When dealing with intractable social problems (e.g. high infant mortality rates), it is unreasonable to expect the intervention of one programme, by itself, to have demonstrable impact. But the programme may well demonstrate significant outcomes. On the basis of these, a programme can legitimately defend itself, as these outcomes may ultimately make a contribution towards achieving an impact at community level. Social problems are recognised as very difficult to change. Evaluations of specific outcomes can therefore help in preventing discouragement if impact is not clearly evident.

Although outcomes and impacts are normally considered as a final stage of the programme and its evaluation, one can monitor outcomes from earlier on in the life of the programme. As pointed out, monitoring service delivery forms the basis of the implementation phase of the programme, but if the programme staff know what outcomes they wish to achieve, they can keep track of them from the beginning. Thus they can develop indicators of change (in knowledge, in behaviour, in functioning) which ought to occur as a result of their intervention. For example, although a score on a standardised reading test may be an imperfect approximation of reading ability, it can serve its purpose as an 'indicator' of that ability. Recording changes in the indicator over time acts as warning system if things are not changing as they should be.

delivery of sexuality information workshops to teenagers, we would need to know how many workshops each teenager actually attended (strength of treatment), and whether the trainers were all following the same curriculum (integrity of treatment). Chapter 7 illustrates a number of instances of how the integrity of treatment was degraded in terms of inconsistency of delivery. Drawing attention to the importance of monitoring the integrity of treatment does not imply that the programme should be fixed and unchanging. Programmes need to adapt to specific local conditions. The point is that these must be documented. If one adaptation is more successful than another, we need to understand exactly how and why this occurred.

These measures of programme implementation also provide feedback on the quality of service delivery. If an assessment can be made quite early in the life of the programme on how well services are being delivered, and problems and difficulties can be identified, such information can be fed back on a very short feedback cycle to improve the programme. This kind of evaluation-of-progress is best used to make adjustments to the delivery of the programme in its early stages.

Whether people are participating in sufficient numbers, and whether the intended recipients are being reached, is another aspect of implementation which needs to be monitored - i.e. the actual population being served (see Posavac & Carey, 1997 above). For example, community workers may be employed to deliver an intervention intended for 15-18-year-old teenagers. When an assessment is done of who actually comes to the events, however, it may turn out that it is mainly 10-12-year-olds (the reason being that community workers found it easier to work with younger children), and were unwittingly arranging activities which appealed more to the younger children). Furthermore, one needs to know, as soon as possible, whether those who participate in the programme are doing well or better.

Finally, simply documenting the progress of the programme, describing its identifiable components (activities, behaviours, products, strategies, materials, etc.), and its planning and delivery, is very useful to others who might want to engage in similar work. It is probably fair to say that in South Africa a vast amount of knowledge about psychosocial interventions goes unused because not enough attention is paid to a careful and detailed description of programmes, their components and processes.

Questions such as 'Did the programme work?', 'Did it achieve its objectives?', 'What were its effects?' are clear indications that there is an interest in the overall outcomes of the intervention. The programme logic outlined above draws our attention to the important distinction between immediate and longer-term effects. Some of the effects achieved by the programme may happen quite soon after the programme has been delivered. Let us call these 'outcomes', after Linney and Wandersman (1991). These are observed or measured changes in the behaviour of target populations. For example, an intervention may change people's knowledge of what causes parasitic infections, or enhance their skills to negotiate with health officials for assistance, or raise awareness in the community of the detrimental effects of parasitic infection on children's lives.

Other effects are further removed in time from the programme, being long-term or

Outcomes and impacts

for the programme recipients. The consequences for the programme, when this kind of outcome monitoring is absent, can be severe.

In the evaluation of the Thousand Schools Project described in Chapter 7, a failure to define and monitor measurable and observable indicators of progress was seen as a major difficulty in the programme not being able to diagnose early problems of implementation. The World Bank has reported that very few impact studies have been done in developing countries, especially of the randomised control nature (Newman, Rawlings & Gertler, 1994). The authors argue that such a neglect handicaps the development community's ability to demonstrate what has been achieved and so to win political support, design more effective projects, and set priorities for resource allocation' (p. 181). The evaluation reported in Chapter 7 is a large-scale impact study that utilised a non-equivalent comparison group design and is therefore a welcome addition to this literature.

Programmes cost money

Is an intervention worth its cost? This is a tough question to ask, and is not often asked of local interventions (Louw, 1998). However, now more than ever before, questions are being asked about the efficiency of psychosocial interventions, and very few interventions would even be considered for funding if not accompanied by cost estimates. Funds are limited, and choices have to be made either to support a particular programme, or to choose between alternatives.

Broadly speaking, one can address the question of programme cost in two ways: one can ask at what cost does the programme achieve what strength of effects (cost-effectiveness); or at what cost does the programme achieve certain savings, expressed in monetary terms (cost-benefit). The difference between the two ways of thinking lies in the way the outcomes are expressed. In the latter, the relationship between costs and outcomes is expressed in monetary terms. In the former, programme outcomes are expressed in terms of substantive improvement or change. The following example illustrates how choices might be made as to which approach to take. When a crèche is established in the workplace for the children of parents who cannot afford to get child-minders, the employer would be interested in what it costs – salaries, supplies, space, electricity, telephones, etc. – as against what it achieves for the children, the staff and the company (cost-effectiveness). These would be difficult to assess. In cost-benefit terms, however, the benefits of the crèche might be more directly expressed in money terms – reductions in staff turnover, less absenteeism, staff loyalty, and length of maternity leave taken.

These are technically difficult questions to answer. They also provide decision-makers with difficult dilemmas to resolve. Decisions have to be made on how to allocate scarce social resources in such a way that communities obtain maximum benefit from them. Some programmes may be more effective than others, in that they achieve more for less money. Chapter 4 refers directly to the costs of the nutrition programme in relation to its effectiveness compared to other programmes aimed at educational improvement. For our society, this is important to know. Indeed, cost-effectiveness and cost-benefit questions may take us right back to the question of priorities. Scarce resources cannot be spent on problems with low priorities.

CONCLUSION

In an attempt to answer the question 'How can evaluation be used to strengthen or improve interventions?', this chapter has presented a way of thinking about programmes and their evaluation. Central to this way of thinking is the notion of a feedback system. In the early stages of the programme, information needs to be fed back to programme staff on how the components of the programme are being delivered: in the intended format, at what strength, and to whom. This relatively short feedback loop is intended to make service delivery more effective. Thus individual psychosocial interventions can be strengthened by receiving practical and timely information about their functioning, and perhaps by meeting the need for a credible judgement of the programme's value.

There are also longer feedback loops involved in interventions and their evaluation, which can also be used to improve them. Programmes can be improved in the long term, through attempts to accumulate knowledge from the many evaluations of individual interventions (see Lipsey, 1997, for a discussion of this). For present purposes, it is sufficient to point out that we can improve our interventions in the long term if we apply, accumulate or synthesise knowledge at every step of the model. At a basic level, for instance, understanding the dynamics involved in the relationships between malnutrition and school achievement, and parasite infection and school achievement, may very well determine initial priorities and choices.

At the level of programme theory, research that goes beyond the immediacy of individual programmes can develop valid, generalisable knowledge about how interventions ought to be implemented, and how the programme processes are supposed to bring about change. Attention has been drawn to programme theory as specifying how the programme is supposed to work. It specifies the change processes that must have taken place in order to link the treatment with the desired outcomes. Thus a causal theory of how the programme works is useful in terms of improving interventions. Understanding and specifying how a particular intervention works gives guidance to the evaluator in a number of ways. It contributes to specifying relevant populations. Understanding and specifying how a particular intervention works gives guidance to the evaluator in a number of ways. It assists in identifying contextual variables likely to be important; it distinguishes between intermediate and ultimate effects; it focuses attention on treatment strength and integrity; it allows stakeholders to understand why and how a treatment will or will not work; and it provides insight into the appropriateness of the programme for the particular social problem.

It ought to be possible to transfer lessons learned from one programme to another. Thus, Schorr (1988) gave a set of lessons derived from successful programmes in poverty relief:

- They offer a broad spectrum of services
- They have organisational flexibility, in that professional and bureaucratic boundaries are crossed regularly
- They are holistic in their approach, in that they see the child in the context of the family and the family in the context of its surroundings
- They contain coherent and easy-to-use services
- They employ committed and results-oriented staff
- They find ways to circumvent professional and bureaucratic limitations to meeting

the needs of clients

- They have comprehensive, responsive and flexible programming

Such a belief (or hope) that lessons can be transferred between programmes underlies this whole book. In Lipsey's (1997, p. 9) words: 'Feedback from cumulated knowledge gained through evaluating thousands of past and present programs should be available in efficient form to help successive generations of programs become progressively better.'

REFERENCES

- Argyris, C. (1982). *Reasoning, learning, and action*. San Francisco: Jossey-Bass.
- Argyris, C. & Schon, D. (1978). *Organizational learning*. Reading, MA: Addison-Wesley.
- Bandura, A. (1977). Self-efficacy: Toward a unified theory of behavioral change. *Psychological Review*, 84, 1185-1193.
- Chen, H.-T. & Rossi, P. (1989). Issues in the theory-driven perspective. *Evaluation and Program Planning*, 12, 299-306.
- Coyle, S.L., Boruch, R.F. & Turner, C.F. (1991). *Evaluating AIDS prevention programs*. Washington, D.C.: National Academy Press.
- de Wilde, J.C. (1967). *Experiences with agricultural development in tropical Africa*. Baltimore, MD: Johns Hopkins University Press.
- Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, Mass.: Addison-Wesley.
- Linney, J.A. & Wandersman, A. (1991). *Prevention Plus III. Assessing alcohol and other drug prevention programs at the school and community level*. Rockville, MD: US Department of Health and Human Services.
- Lipsey, M. (1993). Theory as method: Small theories of treatments. *New Directions for Program Evaluation*, 57, 5-38.
- Lipsey, M.W. (1997). What can you build with thousands of bricks? Musings on the cumulation of knowledge in program evaluation. *New Directions for Evaluation*, 76, 7-23.
- Lipsey, M., Wilson, D.B., Shayne, M., Derzon, J. & Newbrough, J. (1996). *Community needs assessment: The challenges of classification and comparison across diverse needs*. Research Report, Vanderbilt Institute of Public Policy Studies, Nashville, TN.
- Louw, J. (1998). Programme evaluation: A structured assessment. In J. Mouton, J. Muller, P. Franks & T. Sono (Eds.), *Theory and method in South African human sciences research: Advances and innovations* (pp. 255-268). Pretoria: HSRC Publishers.
- Newcomer, K.E., Hatry, H.P. & Wholey, J.S. (1994). Introduction. In J.S. Wholey, H.P. Hatry & K.E. Newcomer (Eds.), *Handbook of program evaluation* (pp. 1-10). San Francisco: Jossey-Bass.
- Newman, J., Rawlings, L. & Gertler, P. (1994). Using randomized control designs in evaluating social sector programs in developing countries. *The World Bank Research Observer*, 9, 181-201.
- Paron, M. Q. (1997). *Utilization-focused evaluation. The new century text* (3rd ed.). Thousand Oaks: Sage.
- Posavac, E.J. & Carey, R.G. (1997). *Program evaluation. Methods and case studies* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
- Schott, L. (1988). *Within our reach. Breaking the cycle of disadvantage*. New York: Doubleday.
- Watson, W.H. & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology*, 49, 156-167.