Health and Social Care in the Community **9**(4), 215–227

Collaboration and partnerships: developing the evidence base

Walid El Ansari[1], Ceri J. Phillips[2] and Marilyn Hammick[1]
[1]School of Health Care, Oxford Brookes University, UK and [2]School of Health Science, University of Wales Swansea, UK

Correspondence
Walid El Ansari
Senior Lecturer in Public Health & Epidemiology
Public and Community Health Department
School of Health Care
Oxford Brookes University
Heritage Gate
Sandringham House
Sandy Lane
Oxford OX4 6LB
UK
E-mail: walid.ansari@brookes.ac.uk

Abstract
Despite the growing literature that collaboration is a 'good' thing, there are calls emphasising the need for evidence of its effectiveness. However, the nature of the evidence to assess effectiveness is less clear. This paper examines the components that contribute to the challenges that confront evidence on collaboration. It considers the differing interpretations that have been placed on evaluation and explores how ways of determining the outcomes of collaboration and the levels of outcome measurement to assess collaborative effectiveness are influenced by the multifactorial nature of the concept. Evidence on the impact of collaboration is influenced by the diversity of perspectives and conceptual facets, and difficulty in measurement of the notions involved. Other factors discussed are the choice of macro or micro evaluation, of proximal or distal indicators, of short and long-term effects, or of individual-level or collective community-level outcomes. The suitability of randomised controlled trials for the measurement of collaborative outcomes as well as the requirement of mixed methods evaluations are highlighted. An evaluation of five community partnerships in South Africa is employed as an example to link the evaluation concepts that are discussed to a real enquiry. If collaboration is to be successful in making a difference in the lives of people, then increasing the precision and context of appraising its effectiveness will reduce the nature of inconclusive evidence and is likely to improve the practice of partnerships, coalitions and joint working in health and social care.

Keywords: collaboration, evaluation, evidence-base, health and social care, methodology, partnerships

Accepted for publication 6 March 2001

Note

For the purpose of this paper, the terms *collaboration*, *partnership*, *coalition* and *joint-working* are used interchangeably. Although some distinctions are made in the literature, in this report, these terms are used to refer to the notion of collective actions by individuals or their organisations for a more shared communal benefit than each could accomplish as an individual player.

Introduction

Numerous recent reports advocate the concept of collaboration within the context of health and welfare services and professional education for healthcare practitioners. The shift is from separatism between services and patronage on the part of the professionals to a partnership model (Statham 2000). In the UK, improved teamwork between professionals is strongly recommended in recent health service publications, e.g. Department of Health (1998). Local partnerships between health service and universities, with a strategic commitment to achieve the tripartite mission of service, teaching and research, have been suggested (Nuffield Trust Working Group on NHS/University Relations 2000). Changes in education and training are fundamental to these arrangements and calls for programmes that are genuinely multiprofessional are currently linked

with workforce planning initiatives (Hargadon & Staniforth 2000). There is also international attention on such collaboration. For example, the Network of Community-Oriented Educational Institutions for Health Sciences supports the interdisciplinary education of the health professionals (NCOEIHS 1991), and in the USA, Latin America and Southern Africa, philanthropic foundations are emphasising partnerships between universities, governments and communities in order to achieve common health goals (WKKF 1992). Similarly, collaborative partnerships between university researchers and their beneficiary communities and target populations have witnessed a forward surge (Grinstead *et al*. 1999, Klein *et al*. 1999, Weiker *et al*. 1999). A partnership is a formal alliance of organisations, groups and agencies that have come together for a common goal (Butterfoss *et al*. 1993), and collaboration is 'to work jointly with others on a project, where those collaborating with others take on specified tasks within the project and share responsibility for its ultimate success' (Michigan State University 1996).

Why is evidence on the agenda?

Evidence-based medicine (Smith 1995, Dean & Hunter 1996) and evidence-based policy (Ham *et al*. 1995) have permeated all spheres of healthcare. Additionally, there is a call for healthcare professionals' education to be engineered on evidence-based partnership approaches (El Ansari & Phillips 1997) and to be more evidential and less anecdotal (Hammick 2000).

Basing policies and practice on the best available evidence is essential to maximising the value of available resources. However, key questions regarding the nature of evidence remain: in addition to evidence of effectiveness and efficiency, evidence related to quality, equity, local ownership and accountability, and political and financial feasibility is fundamental to enhancing outcomes.

The value and importance of evidence in the practice of healthcare is recognised in the work of the Cochrane Collaboration, and the increasing number of teaching programmes on how to conduct evidence-based practice and critical appraisal skills. Taken together, the dual demands of collaboration and evidence are becoming an integral part of many health and social care programmes.

Magnitude of the problem

As partnerships have been widely employed in the healthcare arena, and have existed at the local, state and national levels for several decades (Herman *et al*. 1993), evidence and monitors of good practice are increasingly required for their effectiveness (Butt & Mirza 1996, Morris 1996). Partnerships have developed around a wide range of issues including maternal and child health (Arkin 1986), abortion rights (Staggenborg 1986), sexually transmitted diseases (Tracy 1985) and alcohol and tobacco use (Centers for Disease Control 1990), in response to inadequate state subsidies for family planning initiatives (Herman *et al*. 1993) and for influencing health professionals' education (El Ansari & Phillips 1997, Lazarus *et al*. 1998, El Ansari & Phillips 2001). Philanthropic foundations, for instance, the Robert Wood Johnson Foundation and the Henry J. Kaiser Family Foundation (Tarlov *et al*. 1987), and the W.K. Kellogg Foundation (El Ansari 1994) are sponsoring local partnerships for community health planning and implementation. Granting agencies have invested hundreds of millions of dollars in coalition development as a health promotion strategy (Shopland 1989, Green & Kreuter 1992, Steckler *et al*. 1992).

Hence, the numbers of funded community health projects that rely on coalitions represent considerable investment of resources (Butterfoss *et al*. 1993) and the concept of professionals working together in a collaborative mode has gained wide support (Rawson 1994, Mackay *et al*. 1995). However, despite the general agreement that partnerships are a good thing, Butterfoss *et al*. (1993) called for a systemisation and understanding of what characteristics lead to producing short and long-term impacts on the communities that the coalitions serve. Kreuter *et al*. (2000) confirmed that despite the substantial investment in collaborative mechanisms, relatively little research has been conducted to explore their effectiveness in changing health status in communities. The same lack of evidence is apparent in relation to interprofessional education (Barr *et al*. 1999). In a recent editorial, more evidence and less rhetoric on collaboration in the context of healthcare was recommended (Zwarenstein & Reeves 2000). We are in agreement.

The nature of evidence in the field of joint working

An examination of the collaboration and joint working literature does not readily clarify what constitutes evidence. The randomised controlled trial (RCT), regarded as the gold standard in medical research (Gomm *et al*. 2000) – particularly for testing the efficacy of drugs and other treatments, does not readily lend itself to testing the effectiveness of collaboration. Whilst it is not impossible to randomise an intervention such as new teamworking arrangements, it is more often the case that this is not feasible. Current systematic reviews and meta-analyses rely heavily on the RCT, and the favouring of evidence found by this research design can lead to the exclusion of findings from other types of studies. Consequently, the Cochrane database of systematic

reviews yields a paucity of findings on collaboration as a health promotion intervention. Zwarenstein & Reeves (2000), reporting on the impact of joint nurse–doctor ward rounds on patient outcomes, concluded that 'working together may be worthwhile, but we know too little to glibly assert that collaboration has a positive value'. Moreover, a review of the published literature frequently privileges the industrialised nations and may be biased in favour of certain linguistic capacities (Gillies 1998).

Roe *et al*. (1997) reviewed the published literature on partnership health promotion projects since 1986 using the search strategy of the Cochrane Collaboration. Of 185 identified studies that fitted their inclusion criteria, 43 reported process or outcome data from evaluations that were not simply descriptive. Sixteen of these were RCTs, 15 with control or comparison groups and 12 with pre- and post-testing of impact. Gillies (1997, 1998) reported behavioural change effects on the populations involved in the interventions ranging from 3 to 20%. Zwarenstein & Reeves (2000) queried whether collaboration mattered, and, if it did, whether it could be improved. They concluded, 'we don't know'.

We continue by discussing how ways of enquiring into the outcomes of collaboration are influenced by the multifactorial nature of this concept. We consider the levels of outcome measurement that can reasonably be achieved from such enquiry. The paper also focuses on what constitutes an outcome, and when it is reasonable to measure the effectiveness of collaborative efforts. As the nature, location and purpose of collaborative activity is diverse, this paper is specifically concerned and draws on an example of grassroots partnerships between multiple stakeholder organisations (the academic institutions, health service providers and beneficiary communities). However, the findings are conceptually applicable to a wide range of collaborative efforts: that of front line service deliverers; jointly managed services; strategic partnerships; 'special' initiatives such as the Health Action Zones (HAZs); and community development projects where multiple agencies are involved. Box 1 summarises the challenges confronting the evidence on collaboration.

Box 1 Challenges confronting the evidence on collaboration

- Diversity of perspectives;
- Multiplicity of conceptual facets;
- Difficulty in measurement of notions;
- Selectivity of macro or micro evaluation;
- Variety of proximal or distal indicators;
- Array of short and long-term effects;
- Assortment of individual-level or collective outcomes;
- Measuring a moving target;
- Suitability of randomised controlled trials;
- Requirement of mixed methods evaluations.

Perspectives of collaboration

The very nature of collaboration poses several evaluation and methodological challenges if indisputable evidence is to guide practice. Research questions and subsequent data analysis need to be guided by published theory and empirical work. Zapka *et al*. (1992) asserted that game theorists, psychologists and political scientists have developed formal theory concerning coalition behaviour independently, with limited effort to synthesising perspectives. Murnighan (1978) recommended that more applied field investigation of actual collaborative efforts be undertaken and suggested that all three perspectives had useful aspects to consider. It was noted that the game theorists' emphasis on a coalition's payoffs (outcome achievement), social psychologists' emphasis on resources and political scientists' emphasis on ideology similarity among coalition members were all appropriate variables to include in studies of coalition behaviour. As a result, the literature on collaboration is widely dispersed between several disciplines including organisational management, health promotion, psychology, public health, sociology and public administration. Besides *coalition* theory, Gottlieb & Burdine (1993) suggested a problem-focused theory approach, where the theory of the *problem* is explored, and social research and theory are used to understand how a specific health problem is socially produced and maintained. They also confirmed the importance of theory of the *intervention*, statements or summaries of what is known about the relative success of different intervention strategies with particular populations. The perspectives of these disciplines all need to be taken into account if evaluation philosophy, concept and vision are 'to improve, not to prove' (WKKF 1994).

The conceptual facets of collaboration

The concept of collaboration is built from multiple interlacing facets and it is this characteristic that presents the challenges for evaluating its effectiveness. Diverse standpoints need to be measured in order to weave a holistic picture. Besides the issues of formalisation (Marrett 1971), characteristics of leadership and membership (Knott 1995, El Ansari 1998a), ownership, clarity and operational understanding (Rogers *et al*. 1993), there are the benefits (Rich 1980, Roberts-DeGennaro 1986, Wandersman & Alderman 1993), costs (Guetzkow 1966) and satisfaction with a collaborative effort (Prestby *et al*. 1990). Further aspects include the organisational and personnel barriers (Giamartino & Wandersman 1983, Butterfoss *et al*. 1993, El Ansari 1998b), skills and training (Florin *et al*. 1992, El Ansari 1998c), communication and flow of information (Andrews 1990, Cohen

et al. 1990), representation (Herman *et al*. 1993), quantity and quality of involvement and decision-making (Zuckerman & Kaluzny, personal communication, El Ansari 1998d).

Measurement challenges

The diverse factors that contribute towards the effectiveness of collaboration are not easy to measure. Quantitative measurement tools are frequently employed but indicators of their validity and reliability are not always published, or inappropriate indicators are employed (Main & Pace 1991). Fewer studies on collaboration include information on the psychometric properties of the research tools (Rogers *et al*. 1993, El Ansari 1999). Furthermore, factors such as the nature of interactions between the collaborating parties and decision-making, as well as clarity of roles and operational understanding, sometimes pose measurement challenges. Thus, measurement is frequently reported in terms of the participants' views. The lack of objective independent measures of effectiveness rather than self-reported perceptions may result in an incomplete picture of the evidence (Butterfoss *et al*. 1993, Rogers *et al*. 1993).

An initial question to address in the search for evidence of effectiveness is whether or not the evidence arose out of asking valid and reliable questions that could be answered by the proposed research design. There is also the issue of the appropriateness of the question type – there are those that ask whether an intervention works and those that enquire about why and how it works. The different questions play distinct and separate roles in assessing outcomes. With a social intervention as complex as collaboration, it is important to know more than whether it works if the evidence is to be of practical use. This type of knowledge is much more likely to be revealed by evaluations that use the 'mechanics of explanation', a key feature of realistic evaluation (Pawson & Tilley 1997, p. 55). In other words, the chemistry needs to be revealed and this can only be done in the field. Thus, there is a place in searching for different types of evidence from studies that use eclectic research designs to reveal the nature of how collaboration works. In the next section we examine the evaluation paradigm and its application to measuring the effectiveness of collaboration.

Approaches to evaluation

The question 'What works?' is a completely legitimate one about social policies and the institutions that support them, but 'what works?' is less often asked about the process of programme evaluation itself (Hennessy 1995). Evaluation, like all scientific enquiries, depends on the legitimacy of its procedures to justify its findings and produce credible conclusions. To have faith in the outcomes of an evaluation, the stakeholders must have prior faith in the evaluation design, data collection, and analysis processes that produce the conclusions (Hennessy 1995).

As well as considering the processes and procedures it is increasingly becoming apparent that evaluation must take cognisance of the context within which the policy or programme is located. The notion of realistic evaluation (Pawson & Tilley 1997) is one that provides a useful perspective for attempts to measure the effectiveness of collaboration. It seeks to develop an explanation (and therefore a theory) about how the policy in question works in particular situations/contexts, by exploring the relationships between *context*, *mechanism* and *outcome*. It asks, how does this intervention make a difference in this particular situation? (rather than just asking 'does it work?') and why it might work over here and not over there.

The conceptual geography of evaluation is complex. One commonly used classification is the formative and summative evaluation (Scriven 1996). Formative evaluation is improvement-oriented (Wholey 1996) and provides information to facilitate positive user-led changes based on an evolutionary perspective of project performance and development (Dehar *et al*. 1993, Øvretveit 1998). Summative evaluation is judgement-oriented (Wholey 1996) and focuses upon outcomes, matching these against stated objectives and may include surveys (Dehar *et al*. 1993). However, summative evaluation may neglect the process of an intervention, leaving insufficient knowledge of why a project succeeded or failed (Pawson & Tilley 1997), and is rarely sufficiently time relevant and conclusive to affect policy or budget decision-making. Therefore, formative evaluation is often used to assist in the much more frequent management decisions (Wholey 1996).

Formative and summative evaluations usually involve the *instrumental* use of results (Leviton & Hughes 1981), when some decision or action follows from the evaluation. In contrast, in the *conceptual* use of findings no decision or action is expected (Rossi & Freeman 1985, p. 288); rather, the findings contribute by increasing the knowledge base by testing theory, reducing uncertainty, or increasing clarity (Patton 1996), or a form of informal enlightenment (Weiss 1990, p. 176).

A formal knowledge-oriented approach includes theory-driven (Chen 1990) or theory-based evaluation (Editor's Note 1996). Here, the thrust is to build an evaluation model or conceptual framework that takes account of assumptions and mechanisms underlying a programme (Chen 1990). The use of programme theory connects the evaluation to social science theory, thus

offering potential for increasing knowledge about how effective programmes work (Patton 1996). Programmes should have explicit programmatic theory as programme improvement efforts are virtually impossible if components of the programme and their interrelationships are not specified in advance (Hennessy 1995).

Macro or micro evaluation?

Addressing the level at which evidence is required is fundamental to the sound design of any evaluation. Three levels of evaluation have been described (WKKF 1998). *Project-level evaluation*, executed by appropriate staff at project level, is the consistent, on going collection and analysis of information for use in decision-making. It is not a stand-alone activity, needs to be an integrated part of the project, and provides information about management and service delivery decisions. It addresses contexts, implementation and outcome variables.

Cluster evaluation, on the other hand, determines how well a collection of projects fulfils the objectivity of systemic change. Here, similar projects in 'clusters' are brought together for policy change, which would not be possible in a single or a series of unrelated projects. It looks to identify common threads and shared themes that, having cross confirmation, take on greater significance. It is undertaken by cluster evaluators or the donor body and informs programming strategies and public policy debates (WKKF 1998).

The macro form is *programme and policy-making evaluation*. This addresses cross-cutting programming and policy questions in order to make decisions about programmed funding and support at local, state, and federal level (WKKF 1998). Taken together, these three levels of evaluation provide multiperspective, multisource, multilevel data.

Proximal or distal indicators?

The appropriate distance at which effectiveness or the outcome of collaboration needs to be measured will depend on the aims of the collaborative initiative. It may be feasible to measure the impact of nurse–physician collaboration in intensive care units on patient outcomes (Baggs *et al*. 1999). However, it may prove controversial to attempt measurement of the outcomes of collaboration between service providers and academic institutions on a beneficiary community in terms of improved health. The outcome in this case is too distal to be measured and a change is unlikely to be identified within a reasonable time span, the usual 5-year duration of project funding. Accordingly, more proximal indicators, e.g. satisfaction, commitment, and involvement are often employed.

Fawcett *et al*. (1995) suggested a logic model that addresses process measures, intermediate outcome measures and distal outcome measures that correspond to stages of partnership activity. Distal outcome measures might include reported behavioural changes as a result of the intervention. Hence, the use of monitoring systems, surveys and satisfaction outcomes, behavioural surveys, community-level indicators, or interviews with key participants are each appropriate as the intervention moves from planning to targeting change to producing intermediate or distal outcomes.

In their review of 19 interprofessional education evaluations in the UK, Barr *et al*. (2000) found that only two evaluations measured the impact of the intervention at community or organisation level, with the remainder focusing on changes in the learners' knowledge and attitude or practice-related behaviour. Butterfoss *et al*. (1993) argued that although intermediary outcomes and activities are important, they are insufficient measures of effectiveness. Changes in health status/systems may occur, but go undetected because they are difficult to evaluate, and practitioners may not be asking the right questions when evaluating collaborative activity (Kreuter *et al*. 2000). Gillies (1998), in an analysis of 46 case studies on partnerships from six Regional WHO Offices, reported that the studies focused on process as outcomes. Accordingly, policy development was a major achievement. Gray (1989) argued for collaboration to be viewed as a process and an outcome. In this way, it can be thought of as an evolving forum for addressing a problem where answers to questions about the chemistry and correlates of effectiveness of collaboration are as important as those that measure whether it works, or not.

Short-term or long-term effects?

A related point is the relationship between the short and long-term effects of a collaborative programme (Linney & Wandersman 1991). Long-term effects often extend from short-term ones, but may also encompass system changes and reforms or developing new linkages (Kagan 1991). Ultimately, the degree of impact that collaborative initiatives have on improvements in the health and social status of the communities they serve needs to be measured, but this is difficult and few studies address it. New methodological tools may be required for such assessments (Butterfoss *et al*. 1993).

The time-span of the effect in any measurement of outcomes of collaboration needs consideration. For instance, the temporal nature of the impact is an important question in relation to whether collaborative interprofessional education for undergraduate students will lead to better professional collaboration many years later in the working careers of the graduates, and

inform their practice over time. Pietroni (1994) has highlighted that such evidence is lacking, and improved collaboration and teamwork may require further evidence before the universal assumption that 'learning together' means 'working together' can, without any doubt, be taken on board.

Individual-level or collective outcomes?

Often the outcomes for assessing the effectiveness of community-based partnerships are measures of individual-level changes (Gillies 1998). Shiell & Hawe (1996) have argued that different indicators are required to capture the collective effects, and merely summing individual outcomes is insufficient. Adapting on Cottrell's (1976) work, Goeppinger & Baglioni (1985) concluded that it is difficult for many people to 'avoid confusing community competence with individual and interpersonal competence'. Community level indicators refer specifically to the functioning of the collectivity as a unit, not the functioning of the component parts. Some packages of indicators to measure the effects of collaborative efforts in community-based health promotion are being developed (Centers for Disease Control and Prevention 1997, Kreuter *et al*. 1996). However, a disproportionate amount of the collaborative effort may be geared to individual-level change rather than the more difficult activities targeting policy or environmental change (Kreuter *et al*. 2000).

Collaboration is a broad issue (El Ansari 2000). In some cases the scope of collaboration widens to embrace a more holistic approach to community development rather than a more limited view of improved health. The *gaze* of a programme often broadens to include multiple interventions, stakeholders and partners (El Ansari & Phillips 2001a). The scope of the intervention and the size of the target group distinguishes between community interventions and interventions in communities (Green & Kreuter 1991). This involvement of multiple organisations undertaking multiple interventions staggered in time poses a challenge for the evaluation of collaborative initiatives (Kreuter *et al*. 2000). Furthermore, in the health promotion field, outcomes expected of collaborative mechanisms tend to be broad and ambitious and driven by the priorities of the funder (Kreuter *et al*. 2000, El Ansari & Phillips 2001b). Evaluation will therefore need to accommodate this multiplicity of aims and the diversity and variety of stakeholder and beneficiary groups (Thomas & Palfrey 1996). Hence, there is need for clarity about which programme component, and for which beneficiary group, the evidence is being collected. As Caplan *et al*. (1992) noted, the choice of outcomes can create dilemmas for coalitions, and the challenge for such initiatives is the development of indicators that are multiperspective (Gillies 1998) and that require sophisticated theories of behaviour that take into account the contextual complexities (Gillies 1997). Such characteristics complicate the evaluation task and make it difficult to attribute changes to any particular intervention or strategy.

Measuring a moving target

While some factors are viewed as obstacles in collaboration, precisely the same factors are cited as benefits by other authors. For example, Fiorino (1988) pointed out that the positive outcomes of collaboration include: greater perceived influence over decision-making, improved communication among parties and access to information, and higher quality solutions than those expected from conventional processes. Conversely, Hagebak (1982) and Allensworth & Patton (1990) cited virtually the same domains as the four types of barriers addressed by every coalition, while Habana-Hafner *et al*. (1989) cautioned about barriers, including the degree of authority, the nature of leadership, decision-making and communication.

The apparent controversy between these two opposites is more imaginary than true. Our interpretation is that both groups of authors were referring to the same issues, but with a single difference: the time point in the life of the collaborative effort when the issue was examined. Viewed chronologically, the two sets of issues raised are today's barriers and obstacles, which, if challenged and crossed, become tomorrow's positive outcomes. Thus, attention to the developmental stages of a collaborative effort is critical. Chronologically from formation, implementation, maintenance and accomplishing goals or outcomes (Butterfoss *et al*. 1993), there is considerable overlap from one stage to another, and distinct key factors are associated with the effectiveness of each stage (Kreuter *et al*. 2000). Questions to address are: when in the life span of the collaborative initiative should evidence be collected and how often does it need to be collected? This concept of multiple time samples of evidence also applies to evidence collected to affirm past outcomes, when these are used for comparison in before-and-after studies.

Randomised controlled trials

In the hierarchy of evidence the randomised controlled trial (RCT) is placed very near to the top. The RCT uses the strongest safeguards against sources of confounding that could be responsible for a measured effect (Gomm 2000). A properly conducted and interpreted RCT is superior to any other method of evidence about cause and effect, and thus the effectiveness of health and social care interventions (Shepperd *et al*. 1997). Gillies

(1998) suggested that if the interest is testing the efficacy of a simplistic intervention, then a RCT design with a quantitative data collection method can be applied, but the object and focus of study will be typically narrow. Community-level initiatives do not lend themselves to such constraining designs.

In situations where it is impractical to use a controlled experiment, natural experiments can be used (Gomm 2000). These are events that happen without being created specifically for research purposes. For both the RCT and natural experiments the researcher has control over the variables investigated. This control is accomplished in the RCT at the beginning of the study before the data are gathered, by only including participants who meet certain stringent criteria. In natural experiments control over the variables is accomplished statistically after the data are collected. Data are fitted into different categories in an attempt to make the subjects in each category identical in all respects except one (Gomm 2000). This requirement produces a challenge for natural experiments attempting to detect the effectiveness of collaboration and being able to attribute effects, with precision, to certain variables under investigation.

The conceptual multiplicity and number of variables that foster collaboration means that a RCT will have to control for a wide variety of factors if it is to measure one facet with precision and attribute it. These experimental designs may be of limited value due to community and group diversity and difficulties in controlling variables (Murray 1995). Additionally, the statistical interactions between the array of variables that intermingle and work together to produce an effective collaborative effort will need to be taken into account. For instance, if participants in a collaborative activity are satisfied, then it is likely that they are also involved and committed. Such interactions between concepts of 'satisfaction', 'involvement' and 'commitment' need to be accounted for in any analysis of effectiveness. Multiple regression with hierarchical entry to address the problems caused by collinearity (Neter & Wasserman 1974), as well as path analyses to help identify the most critical ingredients responsible for a measured effect, might be appropriate in such settings.

Finally, if causality is to be inferred, then a temporal relation between collaboration and positive outcomes might not be sufficient to prove the causation. Kreuter *et al*. (2000), reporting on the marginal evidence that coalition strategies and collaborative mechanisms lead to a health status/system change, concluded that these changes might go undetected, because it is difficult to evaluate and demonstrate a cause-and-effect relationship. Furthermore, collaboration may be part of broader interventions that may be partly responsible for an observed health status or systems change.

Thus, while randomised experiments may be the purported gold standard, the challenges for RCTs and natural experiments in the measurement of collaboration effectiveness are numerous. When RCT approaches are not possible, quasi-experimental designs that are high in internal validity could be proposed (Hennessy 1995). These could be either time series (Horn & Heerboth 1982), or multiple comparison group quasi-experimental alternatives to classical randomised experiments. Wholey (1996) further suggested the use of valid performance indicators that could provide time series data on programme outcomes, and challenged evaluators to identify appropriate opportunities to use qualitative evaluation to explore factors contributing to performance evaluations over time.

The role of observational studies also needs to be considered. Recent work to assess the difference in treatment effects shown in observational studies, compared with RCT, indicates that the former can provide valid results and that the previously thought unreliability surrounding observational studies is thrown into doubt (Benson & Hartz 2000). Concato *et al*. (2000) confirmed that both observational studies and the RCT have a role to play in providing evidence for healthcare practice, and the view that the former are always misleading should be revised. These authors demonstrate how recent observational studies using more sophisticated investigation methods can make a sound contribution to the evidence base. There are lessons here for researchers assessing the impact of collaboration. They need a wide awareness of the research methods available to them and the role that each of these can play in any evaluation, if conducted with rigour. Thus, the place of the RCT as the standard for *all* healthcare research and the potential contribution of observational studies to the evidence base may require revisiting.

Mixed/pluralistic methods

The many facets of collaboration dictate that enquiries that aim to dissect and isolate its components are far from simple. It has also been acknowledged that there is no single 'best' approach to adopt in undertaking an evaluation and that there is no one criterion against which one should judge a policy or a programme. Rather, there needs to be consideration given to the range of potential designs, the different types and sources of data that can be accessed and the variety of criteria available for purposes of assessment and the potential trade-offs between them (Phillips *et al*. 1994). Such pluralistic evaluation accumulates evidence from a variety of different sources and employs different research methods in order to generate conclusions concerning outcomes of a project (Billings 2000). A portfolio

of evidence (Beattie 1995) that details the processes and events that take place during and as a consequence of a project is thus compiled.

While quantitative research methods such as the survey will statistically answer questions like 'who is doing what and where', qualitative methods, for example, participant observation studies, focus groups and interviews, are required to describe the change process and answer the 'why and how collaboration works in certain situations' explorative enquiries. A whole system approach, comprising a two-pronged quantitative and qualitative strategy, has been advocated (Popay & Williams 1998), and Gillies (1998) has reported that quantitative methods alone are insufficient for evaluation of community-based collaborative initiatives. Combining both methods is not simply an additive process, but can clearly bring new insights for the study of joint working efforts (El Ansari 1998a, 1999). Milburn *et al*. (1995) warned that the use of mixed methods may produce contradictory results. One trade-off to be considered is between the appropriateness of method for purpose and the rigour of the tool itself (conceptual payoff vs. instrumental payoff). A vital question to address at the beginning of any enquiry into collaborative work is whether the research design and tools are the appropriate ones for the nature of the enquiry and the type of collaboration under investigation. A positive feature of the current work on qualitative systematic reviews is the potential this has for all sound evaluations of collaboration to contribute to the overall evidence base, from a diverse range of enquiry methods.

Logics of change

Partnership work frequently aims at some sort of change or adoption of a new behaviour. In adopting a new behaviour people move through a series of stages (Prochaska & DiClemente 1986): from *precontemplation*, where the change is not thought about, to *contemplation*, where the intention of the change is present, albeit not in the immediate future. The *preparation* stage then follows, where the individual intends to make a change in the immediate future and maybe making small preparatory changes. Then the *action* stage, where attempts at change are active, and finally the *maintenance* stage where the change is adopted and continued, but requires active or conscious effort to be sustained. In community development initiatives, problems arise due to the different rationalities, or understandings and valuing of life processes that exist between communities and health promotion agencies or research teams (Dixon & Sindall 1994). Intertwined is also the question of the desirability of certain sorts of change and the place of changes in the causal chain leading to ultimate improvements in health outcomes. Farquar *et al*. (1985) distinguished between endogenous change (that which occurs with little or no external influence) and exogenous change (that which is precipitated by an external agency), but both change processes frequently intersect creating a range of hybrids. Evaluation problems are pronounced in the partnership situations: the use of standards and indicators that account for shifts in individual behaviour to judge change in community dynamics is inappropriate (Dixon & Sindall 1994). Cappon (1991) acknowledged that in contrast to a matrix of physical and mental indicators, social indicators are the most speculative. Standards and indicators are easy to administer but they reduce complexity overly much and require constant reworking to remain relevant. Here, Dixon & Sindall (1994) suggested that 'Community Stories' and ethnography might be better evaluative tools.

Who should conduct evaluations of collaboration?

External evaluators, who are contracted from an outside agency, have broad evaluation expertise and, because they maintain their positions with their own organisations, they generally have access to more resources than internal evaluators (WKKF 1998). However, this lack of affiliation may have the drawback of detachment from the daily operations of the projects, or limited knowledge of projects needs and goals. Internal evaluators, on the other hand, work within the project, are more familiar with its staff and community members, have access to organisational resources and have more opportunities for informal feedback. However, they may lack outside perspective and technical skills. The role of an effective evaluator includes ability to listen, negotiate, bring together multiple perspectives and analyses of the specific situation and assist in collecting the required evidence (WKKF 1998).

Whoever the evaluation is commissioned to, sound evaluations are grounded in clear and appropriate values (principles, attributes, or qualities held to be intrinsically good, desirable and important) and criteria (standards on which to base judgements) (Stufflebeam 2001). These societal values include: equity (fair to all); effectiveness (successful in meeting targeted needs and/or achieving goals); excellence (possessing high standards); lawfulness (abiding by laws); and citizenship (acting responsibly to the welfare of one's community). The evaluation needs to be built on merit (intrinsic value concerned with whether a programme matches the state of the art) and worth (extrinsic value concerned with meeting the assessed needs of defined beneficiaries). Personnel involved in evaluations are required to have

professional competence (obligations associated with membership in a profession) and job performance (fulfilment of assigned job responsibilities). Evaluation utility standards should ensure that the evaluation will serve the information needs of the intended users. These include: stakeholder identification; evaluator credibility; information scope and selection to address pertinent questions about the programme; values identification so that the bases for value judgements are clear; report clarity, timeliness and dissemination; and evaluation impact to encourage follow-through by stakeholders, so that the likelihood that the evaluation will be used is increased (American Evaluation Association 2001). Similarly, feasibility standards should make sure that the evaluation will be realistic, prudent and diplomatic. Hence, procedures should be practical, to minimise disruption; politically viable and sensitive to the different positions of various interest groups so that their cooperation may be obtained, and any bias to the results can be averted; and cost effective where information of sufficient value can be efficiently produced, so that the resources expended can be justified (American Evaluation Association 2001).

A question to address is who is implementing the evaluation and on whose behalf are they doing so? Many collaborative initiatives are community based with the lay community or voluntary nongovernmental organisations being an integral partner of the effort. Lay involvement in evaluation of their programmes has been widely advocated (Harper & Carver 1999, Sanstad *et al*. 1999, Weiker *et al*. 1999). Although involvement of university/professional researchers in the evaluation is common, the participation of beneficiary community members in the methods employed for collecting and analysing the evidence is not. Here, empowerment evaluation, an approach wholly attentive to the empowering processes, which is based on the use of evaluation concepts, techniques and findings to facilitate improvement and self-determination, might prove appropriate (Fetterman *et al*. 1996). Although the benefits include gaining insights into job-specific tasks and interagency collaboration, as well as acquiring broader professional skills, the stresses involved are the tensions surrounding the evaluation goals and objectives, time commitments and the differing perceptions about the appropriateness of the evaluation methods (Rockwell & Buck 1995).

Multi-stakeholder perspectives

The complexity of collaboration means that its enquiries will exhibit similar features and will have different meanings for each and every participant group. Each constituency will want to ask different questions about whether, how and why it works. Collectively, these may exhibit a slight misfit, with each group seeking evidence for one aspect of the programme. Collaboration is complex and enquiries into its effectiveness by different parties will be on the basis of different agendas with contrasting criteria and potentially conflicting perceptions. The operational and day-to-day features of a programme might be important to the lay community, with policy questions appealing to donors who seek answers across multiple funded sites. Hence, the questions need to be relevant to all participating groups, otherwise considerable confusion as well as frustration may ensue. An outcomes model that takes into account these multiple perspectives has been developed by Barr *et al*. (1999) to assess the effectiveness of collaborative interprofessional education. This enables measurement of a range of outcomes, from the learners' reactions to the intervention to the impact this has had on the community and organisation. The model provides a sound guide to researchers in formulating their study questions so that these lead to a width of outcome measures that will satisfy the different stakeholder constituencies.

Exploring the pervasions of evaluating community partnerships

In order to apply the evaluation concepts discussed above to a real enquiry, an example of an evaluation undertaken in South Africa is outlined in Table 1. During 1996 and 1997, the first author (WEA) evaluated five community partnerships in South Africa (as part of his doctoral work). Several months were spent with each of the five partnerships that participated in the study and data were collected employing a pluralistic mixed methods approach. Table 1 reviews the various evaluation notions and challenges that confront evidence on the effectiveness of partnership efforts, and links them to 'on the ground' activities that were undertaken in the field.

Conclusion

Clearly the challenges confronting the evidence on collaboration can be overwhelming. The diversity of perspectives, multiplicity of conceptual facets, and difficulty in measurement of the notions involved can pose methodological difficulties. In parallel, the choices of macro or micro evaluation, of proximal or distal indicators, of short and long-term effects, or of individual-level or collective community-level outcomes might create technical snags. In addition, the importance of context cannot be minimised in undertaking evaluations, especially of collaboration, partnerships and schemes to develop joint working. Hence, the value of mixed-methods investigations and observational studies need to be highlighted.

Table 1 Evaluation concepts and their application to the South African community partnerships enquiry

Evaluation concept	Enquiry application
1. Theory of the problem	History of South Africa and context of the problem were first studied
2. Participatory evaluation	During the year prior to the study participating partnerships were consulted for their views relating to the research questions
3. Theory of intervention	Published literature was consulted for pertinent theories on partnership functioning
4. Theory-based evaluation	Research tools were complied to reflect synthesis of published theories
5. Realistic evaluation	Several months were spent with each of the five partnerships to gain familiarity with their situation: context, mechanisms and outcomes
6. Evaluation utility standard	Ensuring that information collected would be relevant to partnerships' needs
7. Evaluation feasibility standard: minimal disruption	Data were collected amidst the day-to-day activities of the partners
8. Objectivity	Researcher was not involved with any of the partnerships
9. External evaluator, minimal threat	Researcher was not involved with the funding body
10. Formative evaluation	Aim of evaluation was to provide information to facilitate positive changes
11. Summative evaluation	Study's information on outcomes was matched against partnerships' stated objectives
12. Stakeholder identification; multistakeholder perspectives	Different viewpoints of stakeholders were taken into consideration; academic training institutions, health service providers and beneficiary communities
13. Pluralistic mixed-method evaluation	Research tools were quantitative (survey of 668 respondents) and qualitative (46 interviews)
14. Information scope/selection	Addressing the multiple components that contribute to successful partnerships
15. Perspectives of collaboration (multiple conceptual facets)	Questionnaire was long; interview questions were diverse
16. Measurement challenges	Each partnership aspect was measured by a range of 2–22 questions to accurately capture the dimension
17. Psychometric properties of research tool	A shorter version of the questionnaire was completed by every sixth respondent (test-retest reliability)
18. Triangulation	Survey and interview data were compared for contrasts/confirmations
19. Sensitivity so that bias to the results can be averted	Equal and neutral attention was given to the various stakeholder groups' concerns
20. Cluster evaluation	Information from the five sites was aggregated for commonalties/contrasts
21. Short-term vs. long-term effects	Partnerships were at end of development phase and early into implementation phase, no *a priori* expectations that salient outcomes would have materialised
22. Proximal vs. distal indicators	Short-term effects were inquired about; *certainty* about long-term effects were queried
23. Intermediary measures	Satisfaction levels, commitment, activity and involvement were reported
24. Multiple viewpoints (political viability)	Comparisons were undertaken between the three stakeholder groups
25. Empowerment evaluation	Lay community members were involved with the researcher in collecting data, understanding what is collected, and using computer software for some analyses
26. Measuring a moving target	The study required slightly more than 2 years; last partnership to be evaluated could have progressed and been implemented more than the first one to be evaluated
27. Limitation	Working single-handed, it was not feasible to collect objective independent data on partnerships' performance
28. Cross-pollination	Movement of researcher between the partnerships facilitated exchange of ideas
29. Dissemination	Full interim analyses were undertaken and feedback was provided whilst in the field

An important point to consider is that collaboration needs to be viewed as a process as well as an outcome. When the centre of attention of effectiveness is focused solely on outcomes, the gains and benefits of the process that multiple partner groups go through in finding common ground and working together is at best underestimated, and at worst forfeited. Frameworks for evaluating and improving collaborative efforts shall require the use of diverse and reliable research tools that incorporate process measures as well as intermediate and distal outcome measures (e.g. Fawcett *et al*. 1995, El Ansari & Phillips, 2001c). Only through the purposeful combination of tools and measures (El Ansari 1999) and by examining the questions that are being left out (Kreuter *et al*. 2000) can the evidence-base of collaboration be enriched and the practice of and partnerships be taken forward.

References

Allensworth D. & Patton W. (1990) Promoting school health through coalition building. *The Eta Sigma Monograph Series* **7**, 1–89.

American Evaluation Association (2001) *The Program Evaluation Standards. Summary of the Standards: Utility Standards*. www.eval.org/EvaluationDocuments/progeval.html Accessed 18 March 2001.

Andrews A. (1990) Interdisciplinary and interorganizational collaboration. In: A. Minahan (Ed.) *Encyclopaedia of Social Work,* 18th edn, pp. 175–188. National Association of Social Workers, Silver Springs, MD.

Anonymous (1996) Editor's Note: The origins of theory-based evaluation. *Evaluation Practice* **17** (2), 169–171.

Arkin E.B. (1986) The healthy mothers, healthy babies coalition: four years of progress. *Public Health Reports* **101** (2), 147–156.

Baggs J.G., Schmitt M.H., Mushlin A.I. *et al*. (1999) Association between nurse–physician collaboration and patient outcomes in three intensive care units. *Critical Care Med* **27**, 1991–1998.

Barr H., Hammick M., Koppel I. & Reeves S. (1999) Evaluating interprofessional education: two systematic reviews for health and social care. *British Educational Research Journal* **25**, 533–544.

Barr H., Freeth D., Hammick M., Koppel I. & Reeves S. (2000) *Evaluating Interprofessional Education: a United Kingdom Review for Health and Social Care*. British Educational Research Association, London.

Beattie A. (1995) Evaluation in community development for health: an opportunity for dialogue. *Health Education Journal* **54**, 465–472.

Benson K. & Hartz A.J. (2000) A comparison of observational studies and randomised controlled trials. *New England Journal of Medicine* **342** (25), 1878–1886.

Billings J.R. (2000) Community development: a critical review of approaches to evaluation. *Journal of Advanced Nursing* **31** (2), 472–480.

Butt J. & Mirza K. (1996) *Social Care and Black Communities*. HMSO, London.

Butterfoss F.D., Goodman R.M. & Wandersman A. (1993) Community coalitions for prevention and health promotion. *Health Education Research: Theory and Practice* **8** (3), 315–330.

Caplan P.A., Lefkowitz B. & Spector L. (1992) Health care consortia: a mechanism for increasing access for the medically indigent. *Henry Ford Hospital Medical Journal* **40** (1/2), 50–55.

Cappon D. (1991) Criteria for healthy urban environments. *Canadian Journal of Public Health* **82**, 249–258.

Centers for Disease Control and Prevention (1990) State coalitions for prevention and control of tobacco use. *Journal of the American Medical Association* **264** (12), 1517.

Centers for Disease Control and Prevention (1997) *Community Indicators for HIV prevention: A Delphi Exercise*. HIV/STD Behavioural Branch of the Centers for Disease Control and Prevention, Atlanta, GA.

Chen H. (1990) *Theory-Driven Evaluations*. Sage, Newbury Park, CA.

Cohen L., Baer N. & Satterwhite P. (1990) *Developing Effective Coalitions: a How-to-Guide for Injury Prevention Professionals*. Paper presented at the US Department of Health, Maternal Child and Adolescent Health, Meeting 8.

Concato J., Shah N. & Horwitz R.I. (2000) Randomised controlled trials, observational studies and the hierarchy of research designs. *New England Journal of Medicine* **342** (25), 1887–1892.

Cottrell L. (1976) The competent community. In: B. Kaplan (Ed.) *Further Explorations in Social Psychiatry*, pp. 196–211. Basic Books, New York.

Dean K. & Hunter D. (1996) New directions for health: towards a knowledge base for public health action. *Social Science and Medicine* **42**, 745–750.

Dehar M.A., Casswell S. & Duignan P. (1993) Formative and process evaluation of health promotion and disease prevention programs. *Evaluation Review* **17**, 4.

Department of Health (1998) *A First Class Service. Quality in the New NHS*. HMSO, London.

Dixon J. & Sindall C. (1994) Applying logics of change to the evaluation of community development in health promotion. *Health Promotion International* **9** (4), 425–433.

El Ansari W. (1994) Community partnerships in health; critical lessons learned from a case study of the Hillbrow primary health care project, Johannesburg, South Africa. MSc Thesis (unpublished). London School of Hygiene and Tropical Medicine, London, UK.

El Ansari W. (1998a) Partnerships in health: the pressing challenges. *Public Health Forum* **2** (3), 5.

El Ansari W. (1998b) Partnerships in health: how is it going to work? *Target* **29** (June), 18.

El Ansari W. (1998c) Partnerships and new ways of learning: a second opinion. *NHS Magazine* **15** (Winter), 21.

El Ansari W. (1998d) Tackling health issues at a neighbourhood level – lessons from South Africa. *Health for All Network News* **November**, 14–15.

El Ansari W. (1999) A study of the characteristics, participant perceptions and predictors of effectiveness in community partnerships in health personnel education: the case of South Africa. Doctoral Thesis, unpublished. University of Wales College, Newport.

El Ansari W. (2000) *Health Improvement Programmes – the Future: Partnerships for Better Health*. Report to The Scarman Trust. The Scarman Trust, Birmingham.

El Ansari W. & Phillips C.J. (1997) Altogether better health personnel education? Findings from three community partnerships in health in South Africa. In: J. Conway, R. Fisher, L. Sheridan-Burns & G. Ryan (Eds) *Research and Development in Problem Based Learning*, Volume 4, pp. 127–134. The Australian Problem Based Learning Network, Newcastle, Australia.

El Ansari W. & Phillips C.J. (2001a) Community Development for a Changing World? Innovative Joint Working in Health Care – A South African Partnership Model. *International Journal of Public Private Partnerships* **3** (1), (in press).

El Ansari W. & Phillips C.J. (2001b) Partnerships, community participation and intersectoral collaboration in South Africa. *Journal of Interprofessional Care* **15** (2), pp. 119–132.

El Ansari W. & Phillips C.J. (2001c) Empowering health care workers in Africa: partnerships in health – beyond the rhetoric towards a model. *Critical Public Health*, in press.

Faculty Working Committee at Michigan State University (1996) *Points of Distinction: A Guidebook for Planning & Evaluating Quality Outreach*. Michigan State University, Michigan.

Farquar J., Maccoby N. & Wood P. (1985) Education and communication studies. In: W. Holland, *et al*. (Eds) *Oxford Textbook of Public Health*, Vol. 3, pp. 14–22. Oxford University Press, Oxford.

Fawcett S.B., Sterling T.D., Paine-Andrews A. *et al*. (1995) *Evaluating community efforts to prevent cardiovascular diseases*. Centres for Disease Control and Prevention, National Centre for Chronic Disease Prevention and Health Promotion, Atlanta, GA.

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

Fiorino D.J. (1988) Regulatory negotiation as a policy process. *Public Administration Review*. **July–August**: 20–28.

Florin P., Chavis D., Wandersman A. & Rich R.C. (1992) Analysis of dynamic psychological systems: methods and applications. In: R. Levine & H. Fitzgerald (Eds) *Analysis of Dynamics Psychological Systems*, pp. 341–369. Plenum, New York.

Giamartino G.A. & Wandersman A. (1983) Organisational climate correlates of viable urban organisations. *American Journal of Community Psychology* **11** (5), 529–541.

Gillies P.A. (1997) *The Effectiveness of Alliances or Partnerships for Health Promotion*. WHO Working Paper. Fourth International Conference on Health Promotion. WHO, Geneva.

Gillies P. (1998) Effectiveness of alliances and partnerships for health promotion. *Health Promotion International* **13** (2), 9–120.

Goeppinger J. & Baglioni A. (1985) Community competence: a positive approach to needs assessment. *American Journal of Community Psychiatry* **13**, 507–523.

Gomm R. (2000) Understanding experimental design. In: R. Gomm & C. Davies (Eds) *Using Evidence in Health and Social Care*, pp. 46–64. Sage, London.

Gomm R., Needham G. & Bullman A. (2000) *Evaluating Research in Health and Social Care*. Sage, London.

Gottlieb N.H. & Burdine J.N. (1993) Editorial. Social science theory in health education: time for a new model? *Health Education Research* **8** (3), 305–312.

Gray B. (1989) *Collaborating. Finding Common Ground for Multiparty Problems*. Jossey-Bass, San Francisco.

Green L. & Kreuter M. (1991) *Health promotion planning. An educational and environmental approach*. Mayfiel, Palo Alto, CA.

Green L. & Kreuter M. (1992) CDC's planned approach to community health as an application of PRECEED and an inspiration for PROCEED. *Health Education* **23**, 140–144.

Grinstead O.A., Zack B. & Faigeles B. (1999) Collaborative research to prevent HIV among male prison inmates and their female partners. *Health Education and Behavior* **26** (2), 225–238.

Guetzkow H. (1966) Relations among organisations. In: R. Bowers (Ed.) *Studies on Behavior in Organisations. A Research Symposium*, pp. 13–44. University of Georgia Press, Athens, Georgia.

Habana-Hafner S., Reed H. B. & Associates (1989) *Partnerships for Community Development: Resources for Practitioners and Trainers*. The University of Massachuesetts, Center for Organisational and Community Development, Amherst, MD.

Hagebak B.R. (1982) *Getting Local Agencies to Cooperate*. University Park Press, Baltimore, MD.

Ham C., Hunter D.J. & Robinson R. (1995) Evidence based policy making. *BMJ* **310**, 71–72.

Hammick M. (2000) Interprofessional education: evidence from the past to guide the future. *Medical Teacher* **22** (5), 472–478.

Hargadon J. & Staniforth M. (2000) *A Health Service of All the Talents: Developing the NHS Workforce*. Department of Health, London.

Harper G.W. & Carver L.J. (1999) 'Out-of-the-mainstream' youth as partners in collaborative research: exploring the benefits and challenges. *Health Education and Behavior* **26** (2), 250–265.

Hennessy M. (1995) What works in program evaluation. *Evaluation Practice* **16** (3), 275–278.

Herman K.A., Wolfson M. & Forster J.L. (1993) The evolution, operation and future of Minnesota SAFPLAN: a coalition for family planning. *Health Education Research: Theory and Practice* **8** (3), 331–344.

Horn W. & Heerboth J. (1982) Single case experimental designs and program evaluation. *Evaluation Review* **6** (3), 403–424.

Kagan S.L. (1991) *United We Stand: Collaboration for Child Care and Elderly Education Services*. Teachers college Press, New York.

Klein D., Williams D. & Witbrodt J. (1999) The collaboration process in HIV prevention and evaluation in an urban American Indian clinic for women. *Health Education and Behavior* **26** (2), 239–249.

Knott J.H. (1995) *Building Sustainable Partnerships*. Paper prepared for the W.K. Kellogg Foundation conference. Building Partnerships: An Agenda for Health Around the World. Miami, USA, March 22, 1995. Miami, USA.

Kreuter M., Lezin N. & Koplan A. (1996) *National Level Assessments of Community Health Promotion Using Indicators of Social Capital*. Report to the WHO/EURO Working Group on Evaluation in Health Promotion. WHO, Copenhagen.

Kreuter M.W., Lezin N.A. & Young L.A. (2000) Evaluating community-based collaborative mechanisms: implications for practitioners. *Health Promotion Practice* **1** (1), 49–63.

Lazarus J., Meservey P.M., Joubert R., Lawrence G., Ngobeni F. & September V. (1998) The South African community partnerships: towards a model for interdisciplinary health personnel education. *Journal of Interprofessional Care* **12** (3), 279–288.

Leviton L.A. & Hughes E.F.X. (1981) Research on the utilisation of evaluations: a review and synthesis. *Evaluation Review* **5** (4), 525–548.

Linney J.A. & Wandersman A. (1991) *Prevention Plus III. Assessing Alcohol and Other Drug prevention Programs at School and Community Level: A Four Step Guide to Useful Program Assessment*. Office of Substance Abuse prevention, Rockville, MD.

Mackay L., Soothill K. & Webb C. (1995) Troubled times: the context for Interprofessional collaboration. In: K. Soothill, L. Mackay & C. Webb (Eds) *Interprofessional Relations in Health Care*, pp. 5–10. Edward Arnold, London.

Main D.S. & Pace W.D. (1991) Measuring health: guidelines for reliability assessment. *Family Medicine* **23** (3), 227–230.

Marrett C. (1971) On the specification of interorganisational dimensions. *Sociology and Social Research* **56**, 83–99.

Milburn K., Fraser E., Secker J. & Pavis S. (1995) Combining methods in health promotion research: some considerations about appropriate use. *Health Education Journal* **142**, 569–575.

Morris J. (1996) *Encounters with Strangers. Feminism and Disability*. Women's Press, London.

Murnighan J.K. (1978) Models of coalition behavior: game theoretic, social psychological, and political perspectives. *Psychological Bulletin* **85** (5), 1130–1153.

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

NCOEIHS (1991) *A Short Description of its Aims and Activities*. The Network of Community-Oriented Educational Institutions for the Health Sciences, Maastricht, the Netherlands.

Neter J. & Wasserman W. (1974) *Applied Linear Statistical Models*. Richard D Irwin, Homewood, IL.

Nuffield Trust Working Group on NHS/University Relations (2000) *University Clinical Partnership Harnessing Clinical and Academic Resources*. The Nuffield Trust, London.

Øvretveit J. (1998) *Evaluating Health Interventions*. Open University Press, Buckingham.

Patton M.Q. (1996) A world larger than formative and summative. *Evaluation Practice* **17** (2), 131–144.

Pawson R. & Tilley N. (1997) *Realistic Evaluation*. Sage, London.

Phillips C.J., Palfrey C. & Thomas P. (1994) *Evaluating Health and Social Care*. Macmillan, Basingstoke.

Pietroni P. (1994) Editorial. *Journal of Interprofessional Care* **8** (3), 245–246.

Popay J. & Williams G. (1998) Partnerships in health: beyond the rhetoric. *Journal of Epidemiology and Community Health* **52**, 410–411.

Prestby J.E., Wandersman A., Florin P.R., Rich R.C. & Chavis D.M. (1990) Benefits, costs, incentive management and participation in voluntary organisations: a means to understanding and promoting empowerment. *American Journal of Community Psychology* **18** (1), 117–149.

Prochaska J.O. & DiClemente C.C. (1986) Toward a comprehensive model of change. In: W. R. Miller & N. Heather (Eds) *Treating Addictive Behaviours: Processes of Change*, pp. 3–27. Plenum Press, New York.

Rawson D. (1994) Models of interprofessional work: likely theories and possibilities. In: A. Leathard (Ed.) *Going Interprofessional: Working Together for Health and Welfare*, pp. 38–63. Routeledge, London.

Rich R.C. (1980) The dynamics of leadership in neighborhood organisations. *Social Science Quarterly* **60** (4), 570–587.

Roberts-DeGennaro M. (1986) Factors contributing to coalition maintenance. *Journal of Sociology and Social Welfare* **13**, 248–264.

Rockwell S.K. & Buck J.S. (1995) An interdisciplinary and interagency evaluation team: benefits for other-discipline specialists. *Evaluation Practice* **16** (3), 239–245.

Roe B., Guiness L. & Rafferty A. (1997) *Evidence on the effectiveness of alliances for health promotion: a systematic literature search and qualitative annotated bibliography*. Report to the Health Education Authority, London.

Rogers T., Howard-Pitney B., Fieghery E.C., Altman D.G., Endres J.M. & Roeseler A.G. (1993) Characteristics and participation perceptions of tobacco control coalitions in California. *Health Education Research: Theory and Practice* **8** (3), 345–357.

Rossi P.H., Freeman H. & E. (1985) *Evaluation: A systematic approach*. Sage, Beverley Hills, CA.

Sanstad K.H., Stall R., Goldstein E., Everett W. & Brousseau R. (1999) Collaborative community research consortium: a model for HIV prevention. *Health Education and Behavior* **26** (2), 171–184.

Scriven M. (1996) Types of evaluation and evaluator. *Evaluation Practice* **17** (2), 151–161.

Shepperd S., Doll H. & Jenkinson C. (1997) Randomized controlled trials. In: C. Jenkinson (Ed.) *Assessment and Evaluation of Health and Medical Care: a Methods Text*, pp. 6–31. Open University Press, Buckingham.

Shiell A. & Hawe P. (1996) Health promotion, community development and the tyranny of individualism. *Health Economics* **5**, 241–247.

Shopland D. (1989) ASSIST project targets cancer mortality. *Chronic Disease Notes and Reports* **2** (1), 3.

Smith R. (1995) The scientific basis of health services. *BMJ* **311**, 961–962.

Staggenborg S. (1986) Coalition work in the pro-choice movement: organisational and environmental opportunities and obstacles. *Social Problems* **33** (5), 374–390.

Statham D. (2000) Guest editorial: partnership between health and social care. *Health and Social Care in the Community* **8** (2), 87–89.

Steckler A., Orville K., Eng E. & Dawson L. (1992) Summary of a formative evaluation of PATCH. *Journal of Health Education* **23** (3), 174–178.

Stufflebeam D.L. (2001) *Evaluation Values and Criteria Checklist*. www.wmich.edu/evalctr/checklists/values_criteria.pdf. Accessed 19 March 2001.

Tarlov A.R., Kehrer B.H., Hall D.P. *et al.* (1987) Foundation work: The health promotion program of the Henry Kaiser Family Foundation. *American Journal of Health Promotion*, **Fall**, 74–80.

Thomas P. & Palfrey C. (1996) Evaluation: stakeholder focused criteria. *Social Policy and Administration* **30** (2), 125–142.

Tracy M.K. (1985) Coalition on sexually transmitted diseases. *North Carolina Medical Journal* **46** (April), 237–239.

Wandersman A. & Alderman J. (1993) Incentives, costs, and barriers for volunteers; a staff perspective on volunteers in one state. *Review of Public Personnel Administration* **Winter**, 67–76.

Weiker R.L., Edgington R. & Kipke M.D. (1999) A collaborative evaluation of a needle exchange program for youth. *Health Education and Behavior* **26** (2), 213–224.

Weiss C.H. (1990) Evaluation for decisions: Is anybody there? Does anybody care? In: M. Alkin (Ed.) *Debates on Evaluation*, pp. 171–184. Jossey-Bass, Newbury Park, CA.

Wholey J.S. (1996) Formative and Summative Evaluation: related issues in performance measurement. *Evaluation Practice* **17** (2), 145–149.

WKKF (1992) *Community Partnerships: Redirecting Health Professions Education Toward Primary Health Care*. W. K. Kellogg Foundation, Battle Creek, Michigan.

WKKF (1994) *Improving Cluster Evaluation Information: Some Areas for Consideration*. W. K. Kellogg. Foundation, Battle Creek, Michigan.

WKKF (1998) *Evaluation Handbook*. W. K. Kellogg Foundation, Battle Creek, Michigan. www.wkkf.org/Publications/evalhdbk/default.htm

Zapka J.G., Marrocco G.R., Lewis B. *et al.* (1992) Interorganisational responses to AIDS: a case study of the Worcester AIDS Consortium. *Health Education Research: Theory and Practice* **7** (1), 31–46.

Zwarenstein M. & Reeves S. (2000) What's so great about collaboration. We need more evidence. *BMJ* **320** (15), 1022–1023.