What constitutes evidence in HIV/AIDS education?

Paul Van de Ven and Peter Aggleton

Abstract

It is clear that many sources of evidence have contributed to our grasp of what does and does not work in HIV/AIDS education. Despite this, there has recently been a distinct move to narrow the evidence of success in this field to experimental and comparative work, with randomized controlled trials positioned as the ‘gold standard’. Here we take up the question of what constitutes evidence in HIV/AIDS education. We explore the social and historical factors which ‘privilege’ certain kinds of evidence above others and question whether there exists but one way of understanding what works best in HIV/AIDS education. We draw expressly upon earlier insights and experience in educational evaluation per se and put a case that evidence gleaned through a range of research methods is more useful than exclusive reliance on experimental and comparative work.

Introduction

In 15 years of responding to the HIV/AIDS epidemic, much has been learned about what does and does not work in HIV/AIDS education. Our knowledge of the characteristics that demonstrate effective and non-effective programming in this area comes from diverse sources. Some of the techniques that have been employed to determine whether a particular form of education works are less rigorous, though not necessarily less appropriate to a given set of circumstances, than others (Aggleton, 1997a). At times when it has been necessary to act with urgency—not uncommon throughout the epidemic—HIV educators have had to draw on personal experience, intuition and educated guesses to get education programmes under way. Over time, HIV educators have accumulated knowledge of what works in HIV/AIDS education through observation, especially observations that have been systematically applied. In other circumstances, assessments against the specific objectives of HIV/AIDS education programmes and demonstration projects have provided informative data on the outcomes of interventions. When there have been acceptable sites, participants and ethical considerations, as well as the necessary time and financial resources, randomized controlled trials (RCTs) and other experimental studies have contributed to our understanding of success in HIV/AIDS education.

Many sources of evidence, each with its corresponding strengths and limitations, have informed HIV/AIDS education and related public health policy. No single source has been sufficient to determine the appropriateness and effectiveness of different strategies. Successful approaches to HIV/AIDS education have been determined by bringing together evidence collected in different ways (Auerbach et al., 1994; Coates et al., 1996; Aggleton, 1997a). This is clearly evident in comprehensive lists of the characteristics of effective programming in HIV/AIDS education such as the compendium produced, for example,
by the Washington State Department of Health (Washington State Department of Health, 1994). This lists some attributes of effective programming that are axiomatic (e.g. target the highest risk populations rather than the easiest to reach or teach, or the politically ‘fashionable’). Other features are based on the experiences and observations of educators (use the familiar language or vernacular of the targeted population, take into account cultural and social context, and ensure that programmes resonate with people’s personal experiences). Some characteristics are designed to facilitate outcomes evaluation (e.g. have clearly defined objectives; specify what changes in behaviours are to be achieved and by whom). At times, the recommended strategies are supported by experimental results [e.g. use ‘natural leaders’ among peers—see the work involving key opinion leaders (Kelly et al., 1991, 1992)].

Despite the fact that many sources of evidence have contributed to success in HIV/AIDS education (as measured through such outcomes as behaviour change or declining HIV infections), there has recently been a distinct move to narrow the evidence of success in this field. One way this has occurred is through effectiveness reviews that have privileged certain research methods rather than assessing the methods in terms of their appropriateness to the particular programme or intervention. There has been in the past few years a growing tendency to embrace experimental and comparative work—particularly RCTs as some sort of ‘gold standard’—as the only appropriate way to evaluate HIV/AIDS interventions. Correspondingly, other sources of evidence have been devalued. This push has come from several quarters, mainly in the UK (Oakley et al., 1995; Oakley and Fullerton, 1996) and in the US (Aral and Peterman, 1996; O’Leary et al., 1997).¹

It is appropriate and timely to consider the question of what constitutes evidence in HIV/AIDS education, to consider the social and historical factors which ‘ privilege’ certain kinds of evidence above others, and to question whether there exists one way of understanding what works best in HIV/AIDS education. By doing so, we hope to explore some important issues concerning the research methods that are of greatest relevance to HIV/AIDS education, and how questions of evidence and research methods relate to the policy and practice of HIV/AIDS education. In this paper we examine these issues and draw expressly upon earlier insights and experience in educational evaluation per se to argue that evidence gleaned through a number of research methods is more useful than an exclusive reliance on experimental work. Our paper is a challenge to a certain prevailing orthodoxy, not a comprehensive review of successful HIV/AIDS education.

Evidence and health promotion

Without doubt, evidence is a basic requirement for rational decisions in all walks of life, in business, politics, law and medicine (O’Rourke, 1997). Nonetheless, there are many different and legitimate ways of understanding some things to be true, and of accumulating evidence of what works. Some people bring to their various professions useful characteristics such as business acumen, political nous, wisdom and even the healing touch. These traits may be enhanced through experience, observation and practice, not only through published data and accounts of controlled studies. So too in the HIV/AIDS arena. Educators and health policy analysts should be encouraged to draw upon the evidence which comes from many sources, some of it intuitive or experiential, some gleaned from systematic observation and some based on experimentation. Evidence and information derived from various means should be valued, carefully examined and filtered to see how applicable they are to a given set of circumstances. Likewise, it is important to acknowledge researcher expertise as well as practitioner expertise. Effective evaluation needs researchers skilled at understanding research methods in order to evaluate ‘success’ appropriately.

The debate surrounding the movement towards evidence-based medicine points to the pitfalls of accepting the results of clinical research trials
Evidence in HIV/AIDS education

as the only ‘true’ kind of medical knowledge (O’Rourke, 1997; Sweet, 1998). Not only does such a view fail to recognize the situated, historically and socially constructed nature of human understandings, it also omits the possibility that such understandings can and do change over time. The published evidence is highly selected, incomplete and often dated. The patients in RCTs are often not representative of the general patient population and trials reported in the literature rarely contain more than 50% of patients initially screened. RCTs usually test only one therapeutic agent, even though many patients with common conditions require multiple drugs. Many journals are still reluctant to publish negative or inconclusive outcomes so there is a publication bias for positive results. Major controlled trials take time to set up, conduct, analyse and report, so that by the time benefit can be ascribed to a particular treatment it may have been superseded by another with apparently greater efficacy and fewer side effects (O’Rourke, 1997).

In one recent study of surgical operations, for example, only 39% of treatment evaluation questions could have been answered by an RCT in an ideal clinical research setting (Solomon and McLeod, 1995). Likewise, in the field of nursing, White (White, 1997) has argued that not all nursing problems are capable of being reduced to a clear issue that can be solved by controlled experimental work and many require artistry to find a solution. White has also pointed to the possibility that researchers applying for funding for non-experimental research may be disadvantaged, and the temptation to use experimental findings to justify restricting choice of interventions for both practitioner and patient. Moreover, there is the problem of the gap between research-based knowledge and its implementation in health-related practice (Berrow et al., 1997).

Each of these problems associated with clinical research trials in medicine has its parallel in, and is equally applicable to, experimental studies in the field of HIV/AIDS education. Apart from difficulties in controlling in a multi-faceted, socially-based intervention, another major problem remains. If knowledge of what works in HIV/AIDS education were to be limited to experimental evidence, one would surely want to ensure that the evidence was not skewed in favour of particular interventions and populations. This is certainly not a tenable proposition. Some interventions lend themselves to experimentation (e.g. teacher-led sex education), others do not (e.g. community-based HIV/AIDS education). Likewise, some populations are relatively amenable to experimental work (e.g. children who attend the same school every day), others not at all (e.g. non-gay identifying men who have sex with other men who visit cruising spots on an irregular basis). It is a perverse notion indeed to admit as evidence only that which can be tailored to the experimental suit. We must be willing to be much more realistic and flexible in our approach to evaluation if we wish to capture that which is most relevant to the frequently complex world in which infections occur and may sometimes be avoided.

Gaining evidence

As indicated above, the position taken here is that evidence in HIV/AIDS education comes, by necessity and by design, from diverse sources. This position is supported by researchers with long-standing experience in HIV/AIDS. As Coates et al. [(Coates et al., 1996), p. 6] have recently remarked, ‘a variety of research methods (including observational studies, quasi-experimental designs, demonstration projects, and evaluations of interventions and societal changes) are needed to illuminate effective HIV prevention strategies’. Such a position finds tentative support among researchers from other social disciplines: ‘The realities of social work...render the use of RCTs difficult to realise routinely. Their cost, and the fairly loose relationship that often exists between our understanding of social problems and the responses we make to them...can mean it is not sensible to place all our evaluative eggs in the RCT basket’ [(Macdonald, 1996), p. 43]. Indeed, one of the strongest present day advocates for RCTs in an earlier time recommended, ‘This is not, of course, to say that the procedure of
randomized controlled evaluation is the only means to reliable knowledge, is sufficient in itself, or is always the right approach’ [(Oakley, 1990), p. 193, emphases in original], a view which we fully endorse.

The problems of affording RCTs the same status in social science as in biomedicine have been outlined numerous times, such as in a recent editorial for Health Education Research (Tones, 1997): random allocation of participants, except on a relatively small scale, is expensive and problematic, and so artificial that generalization to ‘normal’ populations is rarely possible. It is almost impossible to stop ‘leakage’ from experimental to control groups or to avoid competing inputs from other sources. Experimental designs may work when comparing a narrowly focussed intervention with no intervention, but educational interventions typically involve complex multifactorial inputs.

Despite the limitations of comparative and experimental designs in health promotion generally, and in HIV/AIDS education specifically, they continue to be enthusiastically embraced in some quarters. The structure of academia and the politics of research funding are contributing agents (Chapman, 1993; Aggleton, 1997b). Academic advancement is largely dependent on research output. Research units rely on grants for their continued existence and individual researchers on funds to generate output. To persons untrained in the social and educational sciences, evidence accumulated through experimental trials has greater face validity. Wherever such persons control the funding of research and they indicate, whether by commission or by omission, that a particular research method is to be employed or is to be given preference in the distribution of research grants, researchers are likely to oblige. In a regimen where, for example, only controlled experimental studies are funded, researchers are faced with the stark choice between abandoning their research endeavours or ensuring that their research proposals include both experimental and control groups.2 The latter option may be pursued irrespective of the merits of the research method or the overall importance of the research question to be explored (Speller et al., 1997).

Community-based HIV/AIDS education campaigns have been highly effective in bringing about rapid, large-scale risk behaviour change, at least initially.3 Bye (Bye, 1990), for example, has described the significant nature of behavioural change among gay men in San Francisco in the early 1980s where community-wide attitudinal and behavioural change was accomplished through planned change in gay community norms, values and institutions. Multiple strategies focussed not on individuals, mass media, and personal testing and counselling, but on full-scale community mobilization, the use of local gay media, interpersonal communication within informal peer networks, grass-roots outreach, building community, and participation and involvement through volunteer work. The effectiveness of the community mobilization strategies was documented through various assessments of the acceptance of the education messages, and high awareness and recall of the messages. A substantial decline in unsafe sexual behaviours among gay men in San Francisco and the decline in the number of new HIV infections among this group provided sound evidence of success. The efficacy of multi-level, community-based interventions in restricting the development of the epidemic among gay men has also been documented in London (King, 1993) and in Sydney (Kippax et al., 1993).

As Bye (Bye, 1990) correctly points out, research methods which might be used in the evaluation of individual-level behavioural change interventions may not be appropriate for the evaluation of community-level strategies that emphasize sociocultural change accomplished through a diffusion process. Carefully designed experiments are rarely possible and certainly not appropriate when the emphasis is on the mobilization of communities, largely informal social networks and peer-to-peer communication. Evidence of success here is best furnished through data on the acceptance and awareness of the educational messages, community change (e.g. the uptake of safer sexual practices),
Evidence in HIV/AIDS education

decreased morbidity (e.g. fewer new infections) and project replication.

For determining whether or not a particular HIV/AIDS education programme has been successful, the standard outcome measures have often been changes in individual’s knowledge, attitudes, beliefs, intentions, behavioural skills (such as condom use skills, refusal skills and negotiation skills) and behaviours (such as needle and syringe sharing, unprotected anal intercourse, and condom use). Self-report remains the most feasible procedure for obtaining information about AIDS risk behaviours in diverse populations, but it is useful to corroborate these data with reliable seroconversion data where possible (Auerbach et al., 1994) or valid population-level HIV incidence rates (Aral and Peterman, 1996). Low incidence and latency periods preclude the use of seroconversion rates as the outcome measure for experimental interventions which usually have few participants and typically are of short duration. However, in situations where HIV testing is normative and frequent (e.g. among gay men in Australia), seroconversion and HIV incidence rates may provide a useful outcome measure of the impact of community and population level HIV/AIDS education.

An alternative approach to understanding the effects of HIV/AIDS education might be to use mathematical modelling to provide evidence for how prevention activity may reduce HIV transmission. Multi-level models incorporating relevant demographic, epidemiological, risk behaviour and prevention parameters can be generated to show how a given educational programme may reduce infection (Auerbach et al., 1994). Furthermore, in a climate of fiscal restraint, cost–benefit analyses setting the cost of the intervention against the medical care and other savings that would be achieved through averting HIV infections could be invaluable for garnering support for a particular intervention. Both of these approaches offer ideal ways of engaging politicians and other decision makers in HIV work. Alongside RCTs, they have a distinctly political tone, for as has recently been suggested: ‘Perhaps the only real justification for the experimental design is ‘political’, and derives from the need to provide ‘hard’ evidence of effectiveness and efficiency in order to gain the funding needed to stay in business!’ ([Tones, 1997], p. ii).

A feeling of déjà vu

Now we want to look back briefly to a pre-AIDS era, and a time when there was a marked shift away from quantitative and experimental evaluation towards alternative, non-experimental evaluation strategies among those working in another area of social intervention. A landmark in the development of educational evaluation was US domestic reaction to the launch of the Soviet Union’s Sputnik I in 1957 and the subsequent launching of a person into space in 1961 (Haney, 1981; Glaser and Silver, 1994; McBride and Schostak, 1995). Questions were raised about the quality and content of American education, prompting federal support for school curriculum improvement in a number of subjects, notably mathematics and science. Large-scale federal and state-funded evaluations of teaching projects were implemented. The majority of these evaluations were quantitative and experimental in nature, and, in retrospect, yielded little by way of helpful information (McBride and Schostak, 1995). Following a great deal of debate in the educational literature throughout the 1960s and early 1970s, a new breed of evaluators and several different prototypes of evaluation gradually emerged. Experimental studies declined from a position as the pre-eminent research method to become by the mid-1970s just one among an array of techniques for gauging success (Jenkins, 1976).

This decline of experimental work in educational evaluation came about for many reasons. There is scope here [drawing on Jenkins, 1976] to mention only a few of the central issues of what was then a complex debate: the tradition of individual measurement associated with experimental studies attracted increasing criticism as being too narrowly focused. Evaluators were encouraged to take a broader view of what is eligible for collection as
evaluation data. Various audiences have different information needs not all of which can be met through traditional experiments. ‘Formative’ evaluations rather than ‘summative’ ones were required to provide the data to improve ongoing courses. Education came to be seen as a dynamic process, subject to adaptation and change, with constant interaction between the end result and the means adopted to achieve it. It was considered bad evaluation practice to undervalue the experience of practitioners, in this case teachers in classrooms. It was recognized that all research is initially based on value judgements of some kind and so no research is totally ‘objective’ or free from bias.

Other key players in the debate argued along similar lines. For example, Cronbach (Cronbach, 1963) called for evaluation that is formative and ongoing, deploying a wide battery of techniques necessary to describe and assess the outcomes of educational programmes. In his view, the greatest service evaluation can perform is to identify aspects of a course where revision is desirable. White (White, 1971) argued that there is no single process of evaluation: there are many different forms of evaluation, all of which should play a part in assessing whether or not a new curriculum is any good. Parlett (Parlett, 1975) criticized the experimental or ‘agricultural–botany’ paradigm of evaluation with its reliance on testing, quantitative data and notions of objectivity, reliability and neutrality. In contrast, a more social-anthropological paradigm with a wider information base, and emphasis on interpreting, explaining and discovering patterns of coherence and inter-connectedness was asked for.

What then were some of the alternative paradigms to emerge in educational evaluation in the 1970s and how might HIV/AIDS education benefit from these earlier understandings of the evaluation process? Jenkins (Jenkins, 1976) summarized six prototypes of evaluation, each serving different purposes and employing different technologies and raising different problems. Here we briefly outline the various prototypes, adapting Jenkins’ (Jenkins, 1976) curriculum evaluation typology to the perspective of the HIV/AIDS educator. Table I sets out the main characteristics, strengths and limitations of each prototype.

Admittedly, Table I oversimplifies the situation particularly in respect of important ways in which the alternative models supplement each other. The key message is that there is a variety of methods for evaluating programmes, serving a wide range of purposes and audiences. The conventional paradigm of educational research (i.e. experimentation) is complemented by other approaches. The objective, goal-based model determines the degree to which desired changes in knowledge, attitudes and behaviour are actually taking place. Goal-free evaluation attempts to escape the ‘tunnel vision’ of goal-based evaluation by assessing the worth of programmes against a set of a priori checkpoints. Self-study or self-deliberation is an ongoing evaluative activity focussing on the process of educational change to gain knowledge critical in strengthening our efforts and those of our clients. Action research, conducted by practitioners and involving the clients or community in the change process, fits within the self-study paradigm. Illuminative evaluation [embraced recently in the broader health education literature (Tones, 1997)] is concerned to portray day-to-day aspects of the setting: to isolate its significant features; delineate cycles of cause and effect; and comprehend relationships between beliefs and practices, and between organizational patterns and the responses of individuals (Parlett and Hamilton, 1972). Decision-making models serve planning decisions at the various stages of programme development: at the outset when setting objectives to suit a particular context; when determining the procedure to follow; when evaluating how well the programme is proceeding; and at the conclusion when deciding whether to reject or recycle the project.

Here we encounter no less than six different prototypes and six different ways of looking at what works in HIV/AIDS education. As has been our thesis throughout, no single model supplants all others. Rather, there are useful ways in which the techniques supplement and complement each
Evidence in HIV/AIDS education

**Table I. Prototypes of evaluation**

<table>
<thead>
<tr>
<th>Prototype</th>
<th>Key emphasis</th>
<th>Purpose</th>
<th>Key activities</th>
<th>Key viewpoint used to delimit study</th>
<th>Outside experts needed</th>
<th>HIV/AIDS educator involvement</th>
<th>Risks</th>
<th>Pay-off</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objectives model</td>
<td>Objectives of the HIV/AIDS programme</td>
<td>To measure target group progress towards objectives</td>
<td>Specify the objectives; measure progress</td>
<td>Programme manager; HIV/AIDS educator</td>
<td>Measurement specialists</td>
<td>Conceptualize objectives; collect and interpret data</td>
<td>Oversimplify goals; ignore processes</td>
<td>Ascertain target group progress towards specific objectives</td>
</tr>
<tr>
<td>Self-study</td>
<td>Staff self-deliberation</td>
<td>To review content and procedures of the HIV/AIDS programme</td>
<td>Discuss HIV/AIDS programme; make professional judgements</td>
<td>Programme manager; HIV/AIDS educator</td>
<td>None, but possible outside authentication or technical help</td>
<td>Committee discussions and decision making</td>
<td>Exhaust staff; ignore values of outsiders</td>
<td>Increase staff responsibility</td>
</tr>
<tr>
<td>Illumination</td>
<td>Description and judgement data</td>
<td>To report the ways different people see the HIV/AIDS programme</td>
<td>Discover what stakeholders want to know; observe; gather opinions</td>
<td>Audience of final report</td>
<td>‘Social anthropologists’</td>
<td>Keep logs; give opinions</td>
<td>Stir up value conflicts; ignore causes</td>
<td>Broad picture of the nature of the HIV/AIDS programme</td>
</tr>
<tr>
<td>Decision making</td>
<td>Decision making and accountability</td>
<td>To facilitate rational and continual decision making</td>
<td>Identify upcoming alternatives; study implications; set up quality control</td>
<td>Administrator; programme manager</td>
<td>Operations analysts</td>
<td>Anticipate decisions, contingencies</td>
<td>Overvalue efficiency; undervalue goals</td>
<td>Programmes sensitive to feedback</td>
</tr>
<tr>
<td>Experimentation</td>
<td>Cause and effect relationships</td>
<td>To seek simple but enduring explanation of what works in HIV/AIDS education</td>
<td>Exercise experimental control and systematic variation</td>
<td>Theorist; HIV/AIDS researcher</td>
<td>HIV/AIDS research; statistical analysts</td>
<td>Tolerate experimental constraints</td>
<td>Artificiality; lack of generalizability</td>
<td>Rules for developing new HIV/AIDS programmes</td>
</tr>
<tr>
<td>Goal-free evaluation</td>
<td>Goal-free checklist</td>
<td>To assess the HIV/AIDS programme’s effects (irrespective of its goals)</td>
<td>Ignore proponent claims; follow checklist</td>
<td>Evaluator</td>
<td>Independent analyst</td>
<td>Make the HIV/AIDS programme accessible</td>
<td>Overvalue documents and record keeping</td>
<td>Data on effects with little risk that the evaluator will be co-opted by proponents</td>
</tr>
</tbody>
</table>

Adapted from Jenkins (1976).

other. In so doing, and by accumulating evidence from a variety of sources, data triangulation may be achieved. If the findings from different sources point in the same direction it is good evidence that a programme really works. So much the better if the data also contain rich information on the context and process so that the intervention may be replicated and improved.
Policy and practice

Given the urgency of the epidemic, and the need for effective prevention and care, it is imperative to examine the important relationship between knowledge and methods, on the one hand, and HIV/AIDS education policy and practice, on the other. The uncompromising position that the only useful data are those derived from RCTs fails to understand the rather loose association between evidence, public health policy and the idiosyncratic, at times contradictory, choices individuals make about their own health. A prime example is that of injecting drug use (IDU) and HIV transmission. There is clear evidence—some from RCTs, though most from comparative and case studies—that multiple prevention measures of education, needle/syringe exchange programmes, methadone maintenance and accepting injecting drug users as legitimate citizens are effective (Wodak, 1996). Some of the strongest evidence comes from data which show declining IDU HIV transmission in countries which have adopted such multiple prevention measures (e.g. Netherlands, Switzerland) compared with those which have not (e.g. US), but see also Stimson et al. (Stimson et al., 1998) concerning recent research in the developing world. Despite this evidence, there is still a reluctance on the part of politicians and policy makers in many countries to rethink their belief in the effectiveness of drug law enforcement and to implement proven measures such as needle/syringe exchanges. The key challenge for HIV prevention in this area is how to establish greater openness among political leaders and decision makers.

Experimental evaluations with their emphasis on behavioural change at the individual level have limited explanatory power, and tend to obscure key social and relational factors involved in behaviour, such as peer pressure, norms, obligations, emotions, cultural beliefs and organizational structures of communities (Auerbach et al., 1994; Aggleton, 1996b). For this reason, research involving ethnographies, social network analyses and community outreach evaluations is needed to illuminate the important linkages between social structure and individual behaviours. For example, it is imperative that we understand the social conditions which may contribute to HIV transmission in developing countries. Lack of access to universal health care, lack of funding for HIV/AIDS education, and inequalities between women and men are some of the many conditions which may fuel an epidemic. Likewise, an understanding of the social, political and cultural dimensions of IDU is needed if we are to deal with vulnerability to HIV among injecting drug users. Broadly, racial and ethnic discrimination and chronic unemployment are two factors that are integral to high rates of IDU and the subsequent transmission of HIV among certain segments of the population in Western industrialized countries. At the group level, initiation practices and notions of trust and friendship have to be understood and taken into account when developing HIV/AIDS education for injecting drug users (Auerbach et al., 1994; Loxley and Ovenden, 1995; Crofts et al., 1996).

One final point about research methods and HIV/AIDS educational practice is that successful interventions are of their essence likely to be multifaceted. In developed world gay communities, for example, where men have adopted safe sex practices and reduced HIV transmission, success has come about not through a single intervention that can be studied experimentally. On the contrary, change has been achieved through multiple approaches which operate synergistically—including community development, skills building, peer education, media campaigns, destigmatization, and counselling and testing (Bye, 1990; King, 1993; Kippax et al., 1993; Auerbach et al., 1994). Similarly, most interventions to reduce the transmission of HIV among injecting drug users contain many potentially active ingredients (Des Jarlais and Friedman, 1996). To gain real insight into successes in gay communities or among injecting drug users, illuminative rather than experimental evaluations are required. An RCT of this element or that may tell us something about an individual educational factor, but it reveals little about its relative contribution as part of a multifaceted
Evidence in HIV/AIDS education

programme. Of even greater concern, it may suggest to those who fund programmes that less expensive, singular interventions are sufficient to achieve programme success when clearly they are not.

Conclusions

We have argued against the hardline position that the only useful evidence in HIV/AIDS education comes from experimentation, including RCTs, and have called for a more encompassing view of what constitutes evidence in this field. The uncompromising proposition that we can only truly know which forms of education are effective after they have been judged against RCT criteria unnecessarily restricts our knowledge base, devalues useful alternative research methods, and compromises policy and practice. Experimental researchers are not the sole repository of knowledge, nor indeed are comparative evaluations the only way in which we have come to know about the world of HIV prevention and care. Much of what we know in HIV/AIDS education depends on insights from practitioners, programme managers, community members, theorists and scholars from non-experimental traditions. Some evidence has been gleaned from true experimental designs, other evidence from qualitative, illuminative and community-based work. Given the limited role that RCTs have in this area, putting a false premium on experimental and controlled studies will weaken rather than strengthen future efforts to minimize HIV transmission.

Looking back to a parallel debate about appropriate research methods for use in evaluating educational programmes more generally, we find that after some early cold war flirtation with experimental comparative designs, educational evaluators came later to adopt a more inclusive stance. Leading educational evaluators from the 1970s onwards have embraced multiple and diverse evaluation strategies. It is recognized that different technologies—goal-based evaluation, staff self-study, experimental control, illuminative evaluation and such-like—contribute worthwhile information according to the various purposes and audiences of the assessment. No single evaluation prototype can answer all relevant questions and none is without its limitations. The strongest evidence is that which is consistent and comes from a variety of investigators, employing dissimilar methods and taking different theoretical perspectives.

The recent focus on experimental studies and the call for more RCTs has one other hidden trap. It diverts our attention from important emerging issues in HIV/AIDS education. At the HIV Prevention Works Satellite Symposium of the XI International Conference on AIDS in Vancouver it was suggested that experimental procedures could contribute little to current priorities in HIV/AIDS education. Other techniques were needed to address priority areas for HIV prevention research, e.g. the effects of multiple interventions; how best to characterize and conceptualize interventions such as peer education, community mobilization and outreach work; how best to involve people living with HIV and AIDS in the design, implementation and evaluation of programmes; and how best to achieve commensurability and synergy between public sector and NGO/community-based responses (Aggleton, 1996b). If the current fashion for more RCTs and other experimental studies persists, there is a real danger that these and other important new directions for HIV/AIDS education will go quite unresearched.

Acknowledgements

This paper was prepared when P. Van de Ven was on a Public Health Travelling Fellowship awarded by the Australian National Health and Medical Research Council. He had discussions with various researchers, policy analysts and practitioners in the Netherlands and the UK but the views herein are the authors’ own.

Notes

1. That said, we recognize that health education and health promotion, especially in the UK, encompasses a variety of positions and approaches. We also recognize that the views
of individual researchers and the official stances of the organizations for which they work change over time.

2. One health promotion manager in England described an interesting approach received from a social research unit (personal communication with first author, 18 August 1997). A group of researchers had secured funding for an RCT of an HIV/AIDS intervention. The problem was that the research group did not have a specific intervention to evaluate and were thus turning to the health promotion manager to provide one. This anecdote draws attention to the absurdity of the methodological tail wagging the health promotion dog.

3. Whereas community-based approaches may have been successful initially, behaviour change in relation to gay men and HIV/AIDS may not have been sustained over time in all countries. This does not mean that the methods used to assess and validate the impact and effectiveness of early prevention measures were inappropriate, merely that there is an important difference between engendering an initial response and sustaining it over time. We recognize also that community-based strategies are usually most relevant to a particular time and place, and to a particular set of historical, social and medical circumstances.

References


Cronbach, L. J. (1963) Course improvement through evaluation. Teachers College Record, 64, 672–683.


Received on November 30, 1997; accepted on September 14, 1998.