Evaluating healthcare education: Issues and methods

Moira Attree *

School of Nursing, Midwifery and Social Work, University of Manchester, Coupland III, Coupland Street, Oxford Road, Manchester M13 9PL, United Kingdom

Summary   This paper analyses the methodological issues inherent in evaluating healthcare education and considers approaches for addressing these.

Recent policies have exhorted practitioners to base their practice on evidence; however in healthcare education the evidence base is not extensive. Whilst educational evaluation has advanced in the last decades, standardised designs and toolkits are not available. Each evaluation has different aims and occurs in specific contexts, thus the design has to fit the circumstances, yet meet the challenge of scientific credibility. Indicators of educational processes and outcomes are not scientifically verified; no toolkit of standardised 'off-the-shelf' valid, reliable and sensitive measures exists. The evidence base of educational practice is largely derived from small-scale, single case studies; the majority of measures are self-devised, unvalidated tools of unproven reliability, thus meta-synthesis is not appropriate and results are not generalisable. Healthcare educational evaluators need valid and reliable assessments of both knowledge acquisition and its application to practice. The need to establish and explain attribution, i.e. the relationship between educational inputs and outcomes is complex and requires experimental/quasi-experimental design. In addition, educational evaluators face the pragmatic challenge of practice in healthcare contexts, where confounding variables are hard to control and resources are scarce.

Introduction

Recent policy imperatives have exhorted practitioners to base their practice on evidence. In healthcare education, however, the evidence base is not

* Tel.: +44 161 275 5631/5333 (Reception); fax: +44 161 275 7566.
    E-mail address: moira.attree@manchester.ac.uk.

extensive and educational practice is not based on cumulative, robust evidence. Whilst educational evaluation has advanced over the last decades, researchers face various methodological challenges, which have endured over various eras or generations of evaluation (Guba and Lincoln, 1989; Rossi and Freeman, 1993; Madaus et al., 2003). Methodological challenges are prevalent throughout contemporary social research and are identifiable in both education and healthcare education research (Kellagan et al., 2003; Madaus et al., 2003). In evaluation research, as in applied research, a potential conflict exists between the requirements of scientific rigor and the need for pragmatism in fieldwork; gaps between evaluation theory and practice are acknowledged (Pawson and Tilley, 1997; Denzin and Lincoln, 1998; Rossi et al., 2004). This paper analyses the methodological issues faced by researchers undertaking educational evaluation within a healthcare context, and considers approaches for addressing these issues.

Evaluation

Evaluation involves the systematic assessment of the nature and worth or merit (Scriven, 1991; Trochim, 2002); its essence is judging and valuing (Scriven, 1991). Rossi et al. (2004, p2) define evaluation as:

'*a social science activity directed at collecting, analysing, interpreting and communicating information about the workings and effectiveness of social programs*'

The aim of evaluation is to assess the design, implementation, management and effectiveness of social programs/interventions, in order to '*judge and improve*' (Rossi and Freeman, 1993, p5). The fundamental goal of evaluation is to provide empirical evidence that aids decision-making (Walden and Baxter, 2001; Pawson, 2002; Rossi et al., 2004).

Epistemological and ontological perspectives

Evaluation is a diverse and evolving discipline (Pawson and Tilley, 1997; Denzin and Lincoln, 1998), with multiple philosophies, models and methods. Various typologies of evaluation theory and method exist (Kellagan et al., 2003; Madaus et al., 2003); classifications vary according to analysts. Two 'classic' theoretical paradigms or meta-theories are identified (Scriven, 2003): positivist approaches follow 'classic' science principles of hypothesis-testing, using quantitative data to test theory, whilst constructivist approaches focus on exploring how people make sense of their experience, using qualitative data to generate theory. A third, more recent addition is the Realist approach (Pawson and Tilley, 1997), which integrates other theories to design an evaluation using mixed methods and data sources. It is based on the assumption that the evaluator already has a theory about *what works, how and under what conditions*.

Guba and Lincoln (1989) classified evaluation according to four generations; categorising the first three as scientific/positivistic and the fourth as naturalistic/responsive, whilst Denzin and Lincoln (1998) identify four major philosophical frameworks: Post-positivism, Pragmatism, Interpretivism and Critical/Normative science. Trochim (2002) also identified four basic theoretical frameworks: Scientific-experimental; Management-systems; Anthropological-qualitative and Participant-oriented, which are broadly comparable to Denzin and Lincoln's (1998) approaches. Scientific-experimental models follow the positivistic paradigm, embracing the values and methods of science and emphasising objectivity and validity; examples include experimental, quasi-experimental methods, economic analysis and theory-driven evaluation (Chen and Rossi, 1992). Management/Systems models originated in business and adopt a systems theory or operational research approach, emphasising comprehensive, organisational context-related approaches: examples include CIPP (Context, Input, Process, Product) (Stufflebeam, 1983). Anthropological/qualitative models also emphasise context, but adopt phenomenological observation principles to retain the importance of human interpretation; examples include Guba and Lincoln's (1989) fourth generation or naturalistic/responsive evaluation. Finally, there are participant-oriented and emancipatory/empowerment approaches, which focus on including participants and stakeholders, such as service-users and clients in evaluations.

However, some approaches, especially those combining mixed methods, do not easily fit into typologies, for example Illuminative Evaluation (Parlett and Hamilton, 1976), Realistic Evaluation (Pawson and Tilley, 1997) and Impact Evaluation (Rossi et al., 2004). Debates regarding the relative merits of each theory/model persist (Kellagan et al., 2003; Madaus et al., 2003); deliberations about whether theoretical approaches can be combined continue. Guba and Lincoln (1989) argued that integrating theoretical approaches constituted *paradigmatic perjury*. More recently, Teddlie and Tashakkori (2003) propose that a third methodological movement, the *Pragmatic Paradigm*, has emerged, that mixes methods to meet evaluation aims. Mixed methods offer the possibility of inte-