Curriculum evaluation in nursing education: a review of the literature

Judith Chavasse BA RGN Dip Nurse Tutors
Postgraduate Student, Departments of Education and Nursing Studies, The Queens
University, Belfast, Northern Ireland

Accepted for publication 29 September 1993

EVALUATION OF NURSING CURRICULA

Evaluation of nursing curricula as a major consideration in nursing education in Britain and Ireland began to be apparent in the late 1970s. It is now recognized as an integral phase of curriculum development, with an increasing number of published reports. Early enquiries into nurse education courses did not generally draw on curriculum evaluation models to structure their study or explain their findings, although there was no lack of reports about nurse training.

In 1977 the World Health Organisation (WHO) published a monograph (Allen 1977) which specifies the following objectives:

1. To enable each individual school [of nursing] to develop an ongoing evaluation project for . . . study, assessment and development.
2. To incorporate a system of evaluation within new demonstration-type nursing programmes.
3. To promote the study of comparative education in nursing.

The following year the erstwhile Joint Board of Clinical Nursing Studies, finding that course planners lacked knowledge and skills to evaluate their courses, produced a package which helped to introduce the practice to nurse educators in Britain (JBCNS 1978).

As a topic, evaluation appears in the nurse education textbooks in the 1980s; Greaves' two volumes on the curriculum progress from a brief description of Stufflebeam's (1971) Context, Input, Process, Product (CIPP) model and reference to curriculum evaluation as being comparable to the evaluation phase of the nursing process (Greaves 1984), to a greatly expanded discussion (Greaves 1987). It is now normal for such texts to include substantial chapters on curriculum evaluation (Gallego 1987, Wells 1987, Kenworthy & Nicklin 1989). Two main themes are manifest which correspond to the related paradigms in general education: for the classical/agricultural/botanical we find the quality assurance model, and for the naturalistic/sociological/anthropological model there is the holistic approach, both drawing on their respective curriculum theorists from general education. Quality assurance is based on either Stufflebeam's CIPP model or the Donabedian triad...
Curriculum evaluation in nursing education

Curriculum type | Evaluation strategy (paradigm)
--- | ---
Objectives model | Classical, quantitative
Process model | Illuminative, qualitative
Combined model | Comprehensive, holistic, quantitative and qualitative

However powerful the influence of holistic nursing models becomes with concomitant process curriculum models, professional validating bodies are likely to seek outcome measures as evidence of nursing competence. Nursing curricula are therefore likely to be developed using a combination of process and outcome models. It is to be expected that the majority of studies evaluating these combined nursing curricula will use a combination of methods.

This assumption is justified on examination of some of the reported studies of nursing curricula. The majority of these use both quantitative and qualitative methodology and draw on multiple data gathering methods.

AIMS AND PURPOSES OF CURRICULUM EVALUATION

These may be classified as follows: validation; meeting objectives; curriculum improvement; evaluation of innovations; understanding the curriculum in whole or in part. None of these is mutually exclusive and for all there is a hierarchy of outcomes from the provision of information, through judging the value or worth or a programme, to making decisions or changing it in some way (Cronbach 1963, Stake 1967, Eisner 1972). In that it necessarily results in a decision, course validation is in a class of its own. Both professional and academic bodies have their own criteria for accreditation with objective rules which must be obeyed. In the past, professional requirements focused on the settings of learning and students' resulting competencies, while higher education bodies looked for intellectual rigour and academic veracity. In some places, there is now evidence of a convergence, both of validation criteria and by the validating agencies.

The systematic model of curriculum evaluation, with its emphasis on whether or not a programme achieves its targets, continues to be influential in the USA (Clark et al. 1983, Blank 1985, Jones et al. 1987, Watson & Herbener 1990). Whether objectives have been achieved is usually a subsidiary issue in the evaluation of European nursing curricula. The evaluation of aims is implied in many studies, but there is seldom explicit matching of learning objectives and outcomes, although the importance of both is acknowledged. It might be expected that the WHO curriculum aims of 'relevant to the health care needs of a . . . country' and 'accountability . . . of
the nurse . . . to her/his client' (Allen 1977) would be addressed, possibly using a qualitative approach, but this has not been demonstrated in publications reviewed.

Curriculum improvement is a theme running through much of the theoretical literature on evaluation and implied in most evaluation reports. Cronbach (1963) suggests that 'the greatest service evaluation can perform is to identify aspects of the course where revision is desirable' and he advocates the need to 'understand how a course produces its effects and what parameters influence its effectiveness'. In the nurse education literature, curriculum improvement as an aim is not often reported, and when it is it seems to be linked with decision-making. Stufflebeam (1971) argues that 'evaluation must always precede the actual making of decisions'. Three reports of decision-making evaluations all use modifications of Stufflebeam's CIPP model (Stufflebeam 1971). Hogg (1990) proposes an adaptation of this model to evaluate a basic nursing curriculum. Two Australian nurse educators also use the CIPP model to evaluate and develop a nursing diploma in the University of New South Wales (Hengstberger-Sims & McMillan 1991), as do Clark et al. (1983) in the USA. Ediger et al. (1983), using the model devised by Stake (1967), assume that all evaluation is for the purpose of decision making.

Innovation

Innovation is a theme explored by Stake (1977), when he discusses the utility of evaluation to course planners and consumers. He describes formative and summative undertaken together as operative evaluation for teachers and administrators, with emphasis on portraying programmes so as to do justice to the unique qualities which are often aspects of new courses. There are various reports of evaluation of innovative nursing programmes, which describe interesting developments (Allen & Murrell 1978, Jarvis & Gibson 1980, Alexander 1983, Di Florio et al. 1989, CroTTY 1990a,b). The aim of these evaluations is to monitor the development of the innovation (formative) and assess its value or success (summative).

The need to understand a curriculum in order to improve it is an aim that recurs in many different contexts, whether the understanding is historical, speculative, ethnographic, aesthetic, phenomenological, hermeneutic, theoretical, normative, critical or evaluative (Short 1991). The examples that will be considered here are historical (O'Connell 1978, Gallego 1983), normative (McClymont 1980) and ethnographic (Dingwall 1977, Treacy 1987, CroTTY 1990a,b).

O'Connell (1978) develops an historical understanding as she explores an innovation commenced in 1948 and follows it through to 1974; the introduction and development of health visitor education in a university. Many of the issues identified and difficulties encountered are common to the experiences of nurse educators as they begin to work in a university setting. Gallego (1983) evaluates basic nursing education in one school of nursing with special emphasis on its history and development, in order to understand the programme which had evolved.

McClymont (1980) takes a normative approach when she seeks to understand the workings of two health visitor courses. She identifies participants' perceptions, the discrepancies between their interpretations, and attempts to reach a consensus view of the ideals of the health visitor education programme. This is a narrow understanding although it provides useful insights. Also seeking to understand the curriculum as a whole but using an ethnographic framework, are Dingwall (1977), evaluating health visitor education, and Treacy (1987), looking at general nursing education.

Understanding aspects of curricula is the aim of another group of nurse evaluators, each seeking in different ways to apprehend curriculum components which are in some way problematic. It is no accident that two concentrate on the teaching and learning of interpersonal skills (Powell 1982, Gott 1984) and three on practice placements (Dean 1985, Jacka & Lewin 1987, Mackenzie 1992), as these are known to be amongst the most complex and sensitive features of nurse education.

CRITERIA FOR AND METHODS OF EVALUATION

Stenhouse (1975) suggests five criteria to be used in evaluating a curriculum: meaning, potential, interest, conditionality and elucidation. To explain meaning, he proposes a philosophical understanding and that one of the objects of an evaluation 'should be to disclose the meaning of the curriculum rather than to assess its worth, though disclosure of meaning naturally invites assessment of worth'. Both Eisner (1975) and Stake (1972) suggest aesthetic appreciation as a mode of apprehending the meaning of curriculum, but Stake (1972) considers that evaluation reports should be available to a wide readership; the description must therefore be simplified to ensure it is generally accessible. In the evaluation of nursing programmes, meaning is recognized as an important issue; here sociology and sociology of education...

Stenhouse's (1975) second criterion is that of potential, for the purposes of the courses: a factor which is pertinent to both the participants' experiences and to their consequent abilities. Interest is an optimistic epithet for problems within the curriculum; problems are said to be interesting if they are manifest in different educational situations or if they have major relevance for a particular educational programme. Although there are differences, there is great similarity between McClymont's (1980) two health visitor courses and Jacka & Lewin's (1987) three cohorts of student nurses; when there are problems, they may be amenable to shared solutions.

Conditionality as a criterion gives due emphasis to the enormous influence exerted on the curriculum by its context. This influence is exerted at every stage: the ethos of the hospital or college and of the profession, availability of resources, approval of the validating body, all of which must be taken into account in planning; the experience of the curriculum as it is lived is conditional upon the dynamics of the student group (Segall & McKay 1984), the quality of the practical placements (McClymont 1980, While 1980, Dean 1985), the expertise and enthusiasm of the teachers, even the resources of the classroom; outcome, especially long-term outcome, is conditional upon the limitations and opportunities of the field in which the ex-participants go to practise. All these local variables are likely to contribute to success or failure, whether this is in relation to specified potentials or in developing insights through the analysis of problems of interests. Cronbach (1963) addresses conditionality as he identifies the importance of understanding 'how a course produces its effects and what parameters influence its effectiveness'.

The last criterion, elucidation, is a scholarly one: whether the evaluation increases understanding of education or contributes to the development of theory. There is relatively little evidence of this criterion in the nursing literature, although Dingwall (1977), Melia (1987), Treacy (1987) and Mackenzie (1992) all provide sociological, but not pedagogical, understandings which could contribute to the development of a sociology of nursing education.

In addition to these five general criteria, Cooper (1976) makes an important distinction between the provision of information and educational decision making. He suggests that the former needs to be clear, objective and unbiased. The latter involves making judgements and is therefore essentially subjective to some extent. Decision making is complex, involving competing values and resources, political issues and practicability. The decision maker has wide responsibilities, not just to the students, the teachers and the institution, but to the wider public. This is especially so in a discipline like nursing where accountability extends to the students' future patients and clients, and must incorporate the exploding knowledge base for health care and the changing needs, values and expectations of patients, within the health services and throughout the entire country (Allen 1977).

Judgement

Stake (1967) states emphatically that 'description is one thing, judgement is another', suggesting that the latter is demanded by teachers and managers but that both are essential for completeness. In his methodology, he provides a matrix in which data about course intentions and what is observed are clearly separated from standards and judgements. His model is applied by Ediger et al. (1983) to evaluating an American nursing programme; they state that objectivity is facilitated because course intentions are matched against observations of each intended outcome.

Amongst his many contributions to curriculum theory, Stake (1972) explores a dilemma for evaluation: completeness versus selection. He examines the values and limitations of analysing the curriculum so as to focus more narrowly, reducing the phenomena under investigation so as to simplify, achieve consensus and perhaps create theory. In contrast, there is the approach of a general description or 'veridical portrayal', by implication a truer, if less defined, picture, opting for what Stake (1972) called 'substantive portrayal . . . to give a recognizable picture of the programme'.

A curriculum theorist contemporary with Stake, Scriven, developed the concepts of formative and summative evaluation, which can be applied to either the whole curriculum or to selected sections (Scriven 1967). In formative evaluation, process is explored; the exercise provides feedback even as the curriculum is being developed and implemented to enable adjustments to be made early in the developmental process. Di Florio et al. (1989) reported on formative evaluation of a course which had been accredited but which was still being developed. Summative evaluation is concerned with appraisal of outcomes, whether these are the curriculum plan prepared for a particular situation or student outcomes as evidence of the success of this plan.

A related classification is the quality assurance model (Donabedian 1969), which uses structure, process and outcome as a typology. Three studies report on the use

Stufflebeam's CIPP model handles the same data in a similar chronology, plus feedback, labelling them context, input, process and product (Stufflebeam 1971, Finch & Bjorkquist 1977). Context evaluation considers whether or not to offer the programme at all and includes its parameters: clients, goals and objectives. Input evaluation relates to the resources and strategies appropriate and necessary to achieve the goals and objectives. Process information is seen by Finch & Bjorkquist (1977) as concerned with determining what objective effects the programme has on the students participating in it, but places little value on the students' subjective experiences or the hidden curriculum. Blank's (1985) evaluation of a community nursing course sought 'to determine the public health nursing content in the preparation of nurses and the degree of consistency that exists among [the] programs'. She makes a basic assumption that professionalism should provide the framework for standardizing the curricula, which seems to have little rational connection with patients' or clients' needs in the Public Health Nursing Service, or indeed to professional accountability as defined by Allen (1977).

**Process**

Process is considered more extensively by Reichardt & Cook (1979). They consider two aspects of process analysis to correspond to two different purposes, to monitor the context, population and programme of the course, providing feedback, and to establish causal explanations within the programme. For Stufflebeam (1971), both process and product evaluation are the outcome measures needed by curriculum decision makers. Product data are here concerned with measuring the exstudents' observable behaviours, assumed to be the programme's effects on its graduates.

These categories of data, whatever terminology is used, can be useful indicators within curriculum evaluation but they are inadequate by themselves. A major omission is the personal experiences of the participants. Without consideration of the participants' degree of satisfaction, of pleasure or displeasure, evaluators are making assumptions that these have no effect on the structure, input, process or product.

**LEARNING OF SKILLS**

Two studies explore the learning of skills (Powell 1982, Gott 1984). Gott (1984) investigated the 'practical nursing skills and social skills' taught to, and needed by, students during and after the introductory course of a general nurse training programme. Powell (1982) sought the opinions of 'student psychiatric nurses' views of their training (as a whole) with particular reference to interpersonal relationships with patients'. Gott (1984) developed hypotheses but was also open to other issues that emerged; she used multiple methods of data gathering and various sources of data. Powell derived all his data
from focused interviews with student nurses and developed several main areas of concern, notably the contradictions experienced by the student nurses.

In contrast, Jacka & Lewin (1987) developed detailed quantification of the context and process of the training general student nurses received in three schools of nursing and in the wards. Although the descriptive findings are now obsolete, the authors state that 'a central aim [is to] ... devise practical procedures for improving clinical learning which can be used ...'.

There is a dearth of outcome studies of basic nurse training which may be associated with either methodological difficulties associated with a population which scatters after the end of its course or with the vast scope of the competencies required of a registered nurse. Outcome evaluation is represented by four studies of post-basic nursing curricula: one of tutors and three of community nurses. Hollingworth (1985) was interested in evaluating whether a Diploma in Nursing Education had prepared nurse tutors to believe in and teach the nursing process. Baseline observations provided evidence that this would entail a change from a medical to a nursing model in their conceptualization of nursing. Evaluation of this change was assessed by questionnaires completed by the tutors after they had been teaching for 6 months. McClmont (1980) sought to evaluate 'the usefulness of the health visitor course as preparation for [that] role as interpreted by tutors and health visitors', as well as to identify differing perceptions of the health visitor role. She explored Stufflebeam's product (exstudents' and tutors' opinions of the usefulness of the course as a whole and in detail) against aspects of his input (whether it met the tutors' goals and the curriculum as depicted in the regulations and syllabus) (Stufflebeam 1971). Cork (1989) looked at health visitors' first year of practice, evaluating their needs after they had completed the course and thus, by implication, what responsibilities they were ready for when they completed their education.

Many of the studies cited draw on a combination of quantitative and qualitative data to obtain and present a naturalistic picture of the course evaluated. Robottom & Jarvis (1987) also used a naturalistic research approach to investigate an innovative curriculum, drawing on phenomenology, ethnomethodology and symbolic interactionism. They explored the integration of health visitor and district nurse students during a single foundation module. Data were gathered by observation, interviews, questionnaires and visits; triangulation was used to create an understanding of whether the integration desired had occurred.

Illuminative evaluation

Illuminative evaluation is an ethnographic approach to understanding curricula; there are two core concepts: the instructional system manifest through the course documentation, and the learning milieu (Parlett & Hamilton 1972). The main working assumption is that a system can only be understood in its wider contexts and that the individual biography of the course is needed because each course is unique; this reflects Stenhouse's conditionality, the influence of context on the curriculum (Stenhouse 1975). In this methodology, the approach is not predetermined; it develops as issues or problems are identified, so can be said to be 'custom built'. The investigation is flexible and a wide range of data gathering methods may be used. The evaluator seeks to recognize themes from which to develop further lines of enquiry, aiming to keep a balance between finding direction and becoming overcommitted to a possible interpretation.

Apart from objectivity, probably the greatest problem with this approach is that of data overload. It may be necessary to shift and refocus so that only a pertinent selection of what has been explored is actually reported. Whiteley (1992) reports on the evaluation of a continuing education programme for newly registered nurses. She gives an example of progressive focusing when, following initial data gathering from the statutory nursing body and the nurses' employers (Health Boards), she is able to clarify the questions to be asked of the actual providers of the programme. This process of refining is crucial to illuminating or making sense of the most significant features of the curriculum.

Alexander (1983) undertook a carefully controlled experiment to test a teaching/learning approach intended to integrate nursing theory and practice. This is a particularly interesting example of the use of qualitative evaluation, in view of the experimental structure with pretest and post-test control group design. Alexander (1983) states that the 'Gestalt concept, that the whole is more than sum of the constituent parts, was basic to the approach adopted ...' and goes on to say that she is using illuminative evaluation. The study as reported corresponds very closely to the illuminative model, except that the researcher was a participant in the project and that progressive narrowing of focus on to selected issues was not manifest in the published report. However, as flexibility of approach is one of the key notes of this model, it seems justifiable to call the study an illuminative evaluation. McMillan (1987) reports formative evaluation of the practical component of a first-level nursing
I. Chavasse

that she as she refers to the 'overwhelming amount of information in order to cross-check the main findings, but it seems that was sometimes difficult to categorize'.

Two other studies using this model are Crotty's (1990a,b) evaluation of a course designed to enable enrolled nurses to 'convert' for registration, and Chamber's (1988) appraisal of a psychiatric nursing programme. Crotty (1990a,b) used a combination of the quality assurance model (structure, process, outcome) with illuminative methodology and process including an element of focusing on topics of particular significance.

CONCLUSION

Most curriculum evaluations in the literature have been reported by nurse evaluators, many of them teachers on the courses evaluated. There is considerable variation in the cogency and scope of the reports, with evidence of increasing rigour in conjunction with a greater use of qualitative methodology. Aims, criteria and methods are drawn chiefly from sociology and/or general education, or sometimes from a management model. This may account for the absence of studies exploring the WHO curriculum aims of relevance to national health care needs and nurses' accountability to their clients (Allen 1977).

The other surprising lacuna is the paucity of outcome studies; this may perhaps be explained by the complexity of the nursing abilities which are the intended product of any nursing curriculum, although sample indicators could be used to create at least a partial profile of curriculum outcomes.

The literature demonstrates that there is much interest in innovatory programmes, in students' experiences and sociological understandings and some concern for specific aspects of the curricula known to be problematic. The increasing number of qualitative or mixed methodology studies is compatible with the value now placed on process curricula combined with the necessity for academic and professional validation.

The purpose of this review is to collate a representative sample of the more important of the evaluations available and, it is hoped, to contribute to the promotion of the WHO objective 'the study of comparative education in nursing' (Allen 1977).

References


Crotty M. (1990b) The perception of students and teachers regarding the introductory module of an enrolled conversion course. Nursing Education Today 10, 366–379.


Eisner E. (1975) Applying educational connoisseurship and criticism to educational settings. In Beyond the Numbers Game
Curriculum evaluation in nursing education

(Hamilton D., McDonald B., King C., Jenkins D. & Parlett M. eds), McCutcheon, Berkeley, pp. 8–22.