A perspective on clinical evaluation in nursing education

M. Star Mahara RN BSN
Instructor, RN/BSN Programme, University College of the Cariboo,
900 McGill Road, Kamloops, BC, Canada V2C 5N3, and MSN Student,
University of British Columbia, Vancouver, Canada

Accepted for publication 30 October 1997

INTRODUCTION

Any institution wanting an overflow audience for a symposium need only plan a program called ‘Evaluation of Nursing Students in the Clinical Area’.

(Woolley 1977 p. 308)

The words spoken 20 years ago by Woolley (1977) are equally as relevant today, and perhaps more so as educators struggle with issues of clinical evaluation from the interpretive and critical perspectives underpinning many of today’s nursing curricula. Clinical evaluation has assumed increased significance in recent years as nurse educators respond to a cry from both education and health care sectors for increased accountability, definitive standards and cost effectiveness. The problems of clinical evaluation puzzle and perturb faculty and permeate discussions at meetings, conferences and workshops. Two issues are particularly problematic. One concerns the objective-subjective debate. Claims around the subjectivity and incomparability of clinical evaluations led educators to search for objective, reliable and valid assessment methods; these were primarily attempts to develop simulation-based measures and to perfect performance checklists and rating scales (Wood 1982, Girot 1993, Clifford 1994). Such evaluation systems, while resulting in more standardized assessments of student performance, fail to address the complex and contextual nature of clinical practice and clinical learning and do not account for the fact that a student’s performance is, in part, a
function of the quality of the learning environment, especially the teacher-student relationship (Benner 1982, McGaghie 1991, De Vore 1993). The second issue centres around the dual role of clinical instructor as teacher and evaluator. This issue is immanently tied up with how barriers to teacher-student partnership in evaluation are maintained through the distinction between formative and summative evaluation.

Two issues in clinical evaluation will be critically analysed in this paper. The objectivity-subjectivity issue is reviewed and the limits of clinical evaluation practices based solely in positivist beliefs are discussed. The teacher-evaluator and formative-summative dualism is challenged on the basis that its primary purpose is to maintain power differentials that impoverish teacher-student relationships. A case is built for evaluation aimed at assessing and maintaining competency levels.

CLINICAL EVALUATION

Evaluation is both a process and a product; the word ‘evaluation’ refers to the process of systematically and objectively determining the merit, worth and value of things, and it also denotes the products of that process (Scriven 1991). It is a process that includes data collection, the subsequent interpretation of the data, and the formation of judgements and conclusions about the meaning of data. Clinical evaluation represents a particular application of the broader discipline of evaluation, specifically a combination of the fields of performance evaluation (evaluation of student work) and personnel evaluation (evaluation of students). Evaluation is described by Scriven as a professional imperative; all professionals are obligated to periodically participate in, and even arrange for, performance evaluations with the aim of being assess and maintain competency levels.

Evaluation of learning is problematic in practice disciplines because it requires the direct observation of students engaged in actual practice in unpredictable clinical environments (Benner 1982, Wood 1986, Ross et al. 1988, Friedman & Mennin 1991). Clinical evaluation has also been complicated by the fact that teachers fulfil multiple and seemingly incompatible roles (mentor, participant-observer and judge/gatekeeper), observation and interpretation of performance are largely subjective matters, and much of the context of performance is outside the control (and often the direct awareness) of both teacher and student. The dual roles of teachers as educators and evaluators has been reconciled by conceptualizing evaluation as also having a dual nature. Efforts to address the challenge of subjectivity and inconsistency in clinical evaluation have been directed toward objectifying clinical practice and standardizing assessment procedures.

Formative and summative evaluation: a false dichotomy?

Clinical evaluation can be seen as having two interrelated functions. The first is an educative one. The evaluation process should provide information to students and teachers on what learning is taking place and what is required in order to improve the teaching-learning process. This is often referred to as the formative purpose of evaluation. The second function is a gatekeeping one; to maintain professional standards and protect the public by assuring that graduates are qualified to practice as autonomous professionals. The twofold aim of this summative or regulatory purpose of evaluation is to make a definitive judgement as to whether a student’s practice meets the standards set by the profession and to determine whether a student has met the academic requirements of a clinical course.

The two functions have traditionally been viewed as separate for several reasons. Explicitly separating teaching from evaluating emphasizes the need to balance the student’s objective of learning and receiving academic credit for that learning, with the regulatory objective of the profession (Schoenhofer & Coffman 1994). Rather than achieving a balance of evaluative focus though, the effect has been to entrench certain differences. For instance, students are encouraged to be involved in the formative evaluation process while only teachers are considered eligible to participate in evaluation decisions of the summative type (Best et al. 1990, Arthur 1995). The responsibility to maintain high professional and academic standards is effectively maintained as something that is not to be shared with students. The assumption is that summative evaluation requires expertise that only resides within the teacher. In addition, summative evaluation is seen as requiring objectivity; something that can be easily compromised by the close teacher-student relationships that develop in clinical teaching (Lenburg & Mitchell 1991). To correct for this source of bias, teachers step out of the subjective mode somewhere near the end of a clinical rotation in order to objectively synthesize all the data that have been collected to arrive at a final judgement about the student’s practice. Because students are viewed as incapable of being objective about their own practice, at least not until they are licensed professionals, their input is limited to formative events.

Some of the fallacies of this sort of thinking should be clear. In reality, teaching and evaluating are inseparable;
each set of practices is intertwined with and dependent
upon the other. Evaluation is an important and useful
teaching strategy, and judgements and conclusions are
being formed about student learning throughout the
teaching process. Evaluation is a continuous and ongoing
process. Formative and summative evaluation are
similarly seamless; ‘there is a formative element in any
summative evaluation’ (Schoenhofer & Coffman 1994 p.
149). The artificial separation of teaching—evaluating and
formative–summative creates a dichotomy that serves to
maintain distance between teacher and student and
perpetuate a belief that students are unable to be ‘knowers’
of quality practice and their own learning. Furthermore, it
is the author’s experience that the two questions: how to
separate teaching from evaluation and whether students
should be involved in summative decisions, are never
issues when working with successful students. These
concerns seem to arise only when working with borderline
and unsuccessful students and so appear to be more issues
of how to deal with the teacher’s negative feelings when a
student’s performance is substandard, and how to counter
the reluctance of teachers to deal definitively with poor
performance (Lankshear 1990, Cohen et al. 1993, Duke
1996).

Evaluation is a threatening process and steps are
required to address this reality or else students cannot
be expected to willingly develop a professional valuing of
the evaluative process (Scriven 1991). An important part
of professional education is identifying one’s evaluation
anxiety and learning how to cope with it effectively.
Professionals who develop this ability are more likely to
value evaluation as a quality assurance mechanism and
participate actively in periodic external reviews of their
practice. Including students in processes for summative
evaluation is an essential teaching strategy both for
learning to judge practice in relation to professional
standards, and for dealing with the discomfort of being
an evaluee.

The issue of objectivity and consistency

Nurse educators are familiar with the story of their
discipline’s struggle for recognition as a profession and
an academic discipline, and the resultant commitment to
positivism and the scientific method (Thompson 1987,
Dickson 1993). A positivist-rationalist perspective frames
evaluation problems as methodological, promotes valuing
of impartiality and certainty, and gives preferential
treatment to quantitative measures and uniform procedures
(Field 1991). For at least the past two decades, the main
challenge to clinical evaluation from this viewpoint has been the subjectivity and inconsistency of clinical
evaluation practices (Woolley 1977, Wood 1982, Bondy
systems and practices have sought objectivity in order to
ensure the fairness, reliability and validity of evaluations.
Objectivity is based on the assumption that the observer and
observed do not influence one another; the evaluator is
able to stand outside the process and objectively capture
and report on the true state of affairs. To increase
objectivity of the observation process, assessment tools
have reduced complex clinical behaviour to manageable
units that can be measured using standardized data
collection instruments (Benner 1982, Friedman & Mennin

Standardized methods

One form of standardized performance exams, the objec-
tive structured clinical examination (OSCE), has under-
gone extensive study (McKnight et al. 1987, Ross et al.
The OSCE is designed to objectively assess clinical skills
performance in simulations of actual clinical situations.
The triple-jump exercise or examination (TJE) is another
popular standardized procedure used as both a teaching-
learning and assessment strategy (Callin & Ciliska 1983,
Chapman et al. 1993, Smith 1993). The TJE uses case
studies to assess a student’s ability to identify potential
and real patient problems and develop a plan of care. It
provides a means for observing how a student handles
new knowledge and formulates the information into a
management or treatment plan. These exams have largely
been successful at increasing comparability of results
through controlling for as many variables as possible, and
creating a judgement-free measurement system. (The
examiner only looks for behaviours to occur, he or she does
not have to interpret.) Subjectivity is effectively reduced
by developing highly structured testing formats and rating
or checklist forms.

Although standardized measuring systems have im-
proved reliability ratings, and are generally perceived by
students and teachers to be more fair and objective than
‘subjective’ methods, their predictive abilities are low
1991). Findings from medical education studies suggest
that performance in one problem or case does not neces-
sarily predict performance on other problems or cases,
particularly in situations that test problem solving and
clinical reasoning ability (Dauphinee 1995). In addition,
several studies have demonstrated a low correlation
between performance scores on simulations and marks
from written and oral tests of knowledge and global ratings
of performance in the clinical setting (Harper et al. 1983,
Roberts & Brown 1990, MacRae et al. 1995). Each method
seems to assess different aspects of clinical performance.

Standardized performance testsuffer other threats to
reliability and validity, primarily because of the high
levels of student anxiety generated, the limited range of
behaviours that can be objectified, a tendency of these
tests to reinforce the rote memorization of skills checklists, and the artificality factor; the more structured the simulated exam, the less it resembles real life practice situations replete with the multiple variables that influence and shape clinical practice (Benner 1982, Friedman & Mennin 1991, McGaghie 1991). Because clinical practice is described in terms of those actions and tasks that can be observed and checked off, standardized assessment methods are limited to technical aspects of clinical practice such as psychomotor, physical assessment, and communication behaviours.

Standardized methods are useful as one option in a comprehensive evaluation system. However, professional practice is multidimensional and requires a combination of approaches to identify and validate competency (Girot 1993). To expand the universe of potential evaluation practices, educators need to explore evaluation from alternate perspectives. Developing credible and useful evaluation methodologies that capitalize on the influence of teacher, student and context on the development of clinical practice and the determination of its worth requires non-traditional paradigms.

CLINICAL EVALUATION AS INQUIRY

Clinical evaluation can rightly be viewed as a type of disciplined inquiry requiring practices and standards for collecting, interpreting and judging information about the overall quality of a learner’s clinical practice. The evaluation of a student’s practice involves the collection of data through various means; the analysis and interpretation of the data according to standards, competencies and values by which nursing practice is judged; the conclusion, judgements, and decisions that are arrived at about a student’s practice; and recommendations as to what action should follow. As a member of the inquiry family, clinical evaluation is the sister of research. The purpose of clinical evaluation as a form of research is the discovery and verification of the process and product of the teaching and learning of nursing practice.

Nursing educators face the dilemma of how to determine a student’s competence level, while at the same time realizing that this cannot be known directly, exactly and for all time because of the indeterminate, evolving and contextual nature of clinical competency (Field 1991, Hepworth 1991). Processes used must be capable of accurately describing and interpreting a student’s current practice and making predictions about a student’s future capabilities. Professional bodies require objective outcome measures to provide some guarantee of a student’s safety to practice. There are bio-physical aspects of nursing practice that rightly demand that control and predictability be addressed in evaluation systems (Wolfer 1993). Evaluation processes must also be able to capture the uniqueness and context-dependent nature of clinical practice and contribute to greater understanding of both clinical practice and clinical teaching-learning. In addition, evaluation is a political process, so that approaches to evaluation should have the potential to lead to emancipation of the evaluative participants as well as stimulate action aimed at transforming undemocratic clinical and education systems (Bevis & Watson 1989, Guba & Lincoln 1989). Because nursing education and practice deal with these different aspects of reality, different ways of knowing have evolved (Jacobs-Kramer & Chinn 1988, Wolfer 1993). Methodologies for assessing and evaluating practice must similarly address these multiple realities.

The debate over congruence between ontology, epistemology and methodology and the mixing of diverse methods has been argued elsewhere (Guba & Lincoln 1989, Ford-Gilboe et al. 1995). Several authors oppose combining practices from different paradigms. Guba and Lincoln are leaders in this regard, believing that although methods are not paradigm specific, the basic assumptions and beliefs of a paradigm shape how one uses and understands any one inquiry strategy or technique. On the other hand, Malek (1988) and Ford-Gilboe et al. argue convincingly for combining data collection and analysis strategies associated with differing paradigms, maintaining that the separation of methods is artificially limiting. Combining multiple methods adds breadth and depth to an investigation, and increases the persuasiveness of the findings. If evidence to support a decision or judgement is provided through diverse methods, the findings are more likely to be credible and not dependent on the method that was used (Hall & Stevens 1991). Thus, decisions about the quality of a student’s clinical practice are more trustworthy when based on a combination of evaluation data obtained from a variety of methods such as: performance checklists, rating scales and written exams; self-evaluation methods that include numerical rating and narrative formats; the various peer evaluation strategies; and observation of and interview/discussions about performance in actual clinical situations as well as in simulated scenarios.

The advance that may prove most useful in addressing the problems of clinical evaluation is the recognition that subjectivity is not incompatible with credible clinical evaluation processes (Hepworth 1991, Van der Vleuten et al. 1991). Indeed, when the teacher functions as the evaluation instrument, the range of clinical situations and number of influencing variables that can be assessed is greatly increased, as is the probability that the evaluation has adequately represented the student’s practice. Therefore, instead of designing methods that bypass the teacher’s knowledge and skill, evaluation systems should capitalize on the teacher’s professional expertise. Expert nurse educators should know what quality clinical practice entails and can judge both the quality and appropriateness of a student’s practice in relation to
individual client situations (Friedman & Mennin 1991, Girot 1993). This ability to accurately examine, discuss and judge nursing practice in a clinical situation requires the teacher to draw upon her or his personal experience and knowledge of nursing, education and evaluation to apprehend what is salient, determine what needs to be examined further, and to judge the soundness of the student’s practice rationales in each particular instance (Guba & Lincoln 1989). Thus, like nursing practice and research, clinical evaluation involves both subjective and objective elements, utilizes multiple ways of knowing and is as much an art as it is a science (Scriven 1991, Sandelowski 1993). However, as with all fields of skilful practice, nurses do not become expert teachers and evaluators overnight. It is imperative that clinical instructors are properly educated and thoughtfully mentored in order to develop this expertise (Wood 1986, Orchard 1992, Clifford 1994, Duke 1996).

Expanding the education and practice of clinical instructors to include evaluative processes based on interpretive and critical paradigms adds to the teacher’s and student’s ability to understand the student’s practice and to make reasoned and meaningful judgements about what it amounts to. Evaluation practices that combine principles of fourth generation evaluation and qualitative data collection and decision making procedures provide the opportunity for a rich, holistic and artistic portrayal of a student’s practice while providing for responsible evaluation through adhering to a disciplined and systematic approach.

Fourth generation evaluation, constructivism and clinical evaluation

The emergence of fourth generation evaluation paralleled the expansion of science and research into the realm of interpretive and critical world-views. According to Guba & Lincoln (1989), the field of evaluation can be seen to have passed through three distinct phases or generations, each developing in response to the perceived limitations of the previous formulation. Early evaluation was strictly technical and based in measurement. The variables to be evaluated were determined and methods developed to gather scores or measures of the variables. The second generation expanded to include description. Objectives for the evaluation guided the type of information to be collected in order to describe how well the objectives had been achieved. In third generation evaluation, the purpose of measurement and description was to inform decisions and judgements about the merit and worth of the evaluation (the object of the evaluation, i.e. the person, programme, product). Fourth generation evaluation evolved from a concern that the previous forms did not fairly and adequately involve stakeholders in the design of an evaluation and implementation of its results, and from a critique of the positivist paradigm as a guiding methodology for the evaluation of what are in effect, human enterprises.

Fourth generation evaluation and qualitative research share a common grounding in interpretive and critical paradigms (Guba & Lincoln 1989, Ford-Gilboe et al. 1995). Interpretive and critical views both fit comfortably beneath the overarching paradigm of constructivism. Fourth generation evaluation has the following central tenets: a belief in the socially constructed nature of reality, an appreciation of how knower and known interact and shape one another, and a recognition of the situational constraints and power relations that shape inquiry. In clinical evaluation, this translates to teachers and students who work jointly to co-create a picture of the student’s practice and then to determine what the practice means in relation to professional standards. The tenets of constructivism require that the authority and power of the teacher is acknowledged and utilized to create a relationship that empowers students to assume increasing responsibility for participating in the description and judgement of their practice. Teachers and students come to the evaluation having differing concerns, issues and knowledge and skill levels. The evaluation process must facilitate the articulation and exploration of these differences so that a more informed and sophisticated understanding of each other’s position and how it may influence the teaching-learning process, the student’s practice and evaluation itself, is possible. Dialogue and negotiation are key processes for establishing and maintaining the process and rules of the evaluative relationship and for maintaining the integrity of the interactions to assure the quality of the evaluation process and product (Guba & Lincoln 1989).

Fourth generation evaluation fits well with the view of evaluation as a method to illuminate and further the understanding of the teaching and learning of nursing practice. The complex, diverse and ambiguous nature of nursing practice and teaching-learning makes the interactive, meaning-making and context-sensitive nature of constructivism attractive as a basis for evaluation. Fourth generation evaluation is also of value because it directs educators and students to problematize the content, processes and institutions of nursing education and health care. A credible evaluation examines the influence of the clinical environment and the teacher’s practices on the student’s performance. Fourth generation evaluation also requires that educators believe that students can be knowers of quality nursing practice and should be involved in summative decisions. Evaluation processes become a teaching-learning strategy for nurturing a student’s commitment to evaluation as a means for ensuring professional competence. Having participated in directing the evaluation, drawing conclusions and recommending a plan of action, the student has good reason to support the suggestions for future practice and areas of evaluation (Guba & Lincoln 1989).
The credibility of fourth generation evaluation

Constructivism as the basis of evaluative practices suffers from the same prejudices as qualitative research in terms of the long-standing debate over the credibility of its methods and products. Professional bodies and nurses in general are wary of process-orientated practices, preferring instead, methods and practices with an outcome orientation. It is critical that criteria and processes by which to carry out evaluation and judge its quality and goodness are clearly outlined and followed. Evaluation within both positivist and constructivist paradigms can be judged by the various criteria established for determining the credibility of research data, interpretations, judgements and outcomes (Sandelowski 1986, Guba & Lincoln 1989, Hall & Stevens 1991).

The quality and goodness of the data and the interpretations and judgements made from the data can be maintained by strategies for attending to issues influencing the individual teacher or student as well as the teacher and student in relationship. The teacher’s subjective responses to students is shaped by many variables, conscious and unconscious, and may positively or negatively impact teaching and evaluation practices (Paterson & Groening 1996). Reflexivity, ‘the ability to critically examine and use previous experience to influence further action’, is facilitated by awareness of the need for regular self-examination and the availability of methods to monitor the influence of personal values, beliefs and assumptions on one’s perceptions and acts (Paterson 1994 p. 309). Methods utilized by qualitative researchers to examine sources of reactivity in researcher–research participant interactions provide a deliberate and thoughtful structure for analysing intra and interpersonal influences on the collection and interpretation of data (Hall & Stevens 1991, Paterson 1994). It is suggested that teachers maintain a log for questioning and analysing their subjective responses to students and how such behaviour is affecting the course of teaching-learning and the evaluation process.

Teacher–student relationships that are honest, consistent, open, respectful and supportive are needed before students will risk the level of self-disclosure and self-critique required in this type of evaluation system (McBride & Skau 1995). Depth and scope of evaluative data requires relationships characterized by engagement. Rapport can be assessed by the specificity and type of information shared and indications of student comfort and willingness to share (Hall & Stevens 1991). Students cannot be expected to support a fourth generation model unless they feel they have the capability and are encouraged to challenge and question the teacher, experience dialogue and negotiation as a consensual meaning-making process and not as a disguised presentation of the teacher’s predetermined interpretation of an event, and see evaluation as generating growth rather than a list of weaknesses and negative critique.

The processes and products of evaluation must also be open to an external review procedure aside from formal appeal systems which create adversarial relationships among the persons involved (Scriven 1991, Hunt 1992, Orchard 1994). A formal committee of teachers and students can be established to monitor the clinical evaluation system and the faculty and students who make up the system for the ability to adhere to principles of partnership, responsibility and accountability and to check for evidence of consistent privileging of any one construction over another. Peer group discussions are another forum for submitting one’s evaluation practices for external review. Colleagues are helpful in posing questions that assist the evaluator to understand her or his own posture, values and role in the inquiry.

CONCLUSION

There is no doubt that the evaluation of clinical competence is an important activity with high stake consequences (McGaghie 1991). Scriven (1991 p. 43) states this position clearly, ‘doing evaluation and doing it well matters in pragmatic terms because bad products and services cost lives and health, destroy the quality of life, and waste the resources of those who cannot afford waste. In ethical terms, evaluation is a key tool in the service of justice, in programs as well as in personnel evaluation’. The outcomes of clinical evaluation decisions have serious implications for students in terms of self-esteem and protection of their rights to pursue a livelihood; for patients and clients with regards to protection from unsafe practice; and for teachers concerned with potential charges of unfair evaluation practices and the legal repercussions of decisions where students are wrongly failed (or passed) in a clinical practice course (Cohen et al. 1993, Orchard 1994). Fourth generation evaluation with its valuing of stakeholder involvement has much to offer educators faced with developing clinical evaluation systems.

Clinical learning is the heart of the educational experience for nursing students and their teachers. Clinical evaluation processes are more than merely one aspect of clinical learning; they are the quintessence. Curricula that are based on the concepts of meaning-making, reflection and teacher-student partnership provide a basis for evaluation approaches which move beyond the concept of measurement and verification of clinical skill levels to the explication and judging of a student’s clinical practice as a teaching-learning strategy. The incorporation of multiple ways of knowing and judging expands the possibilities for describing and verifying the outcomes of clinical education. The move to multiple paradigms in some schools of
nursing may be what is needed to end what Woolley (1977) termed the long and tortured history of clinical evaluation.

Acknowledgements

The author wishes to acknowledge Dr Barbara Paterson for her mentorship, support and encouragement and her editorial assistance. Thanks also to colleagues in the clinical teaching course at the University of British Columbia for their caring and critical responses to earlier versions of the ideas presented in this paper.

References


