QUALITY ASSESSMENT OF DOCTORAL PROGRAMMES

Shaké Ketefian
EdD, RN, FAAN
Professor; Director, Doctoral/Postdoctoral Studies And International Affairs
University of Michigan, School of Nursing

ABSTRACT

The purpose of this article is to focus attention on the different ways in which doctoral education can be evaluated. It reviews highlights from the literature in this regard, identifies purposes of doctoral education, and different assessment approaches that have been used. The author develops a practical framework to guide educators in their thinking about evaluation. The practical framework is then submitted to a test of its validity through a discussion of criteria and procedures used in one setting. Compelling reasons are provided for the benefits of conducting doctoral programme assessment – for the institution, for society and for other constituents.

QUALITY ASSESSMENT OF DOCTORAL PROGRAMMES

Quality assessment in doctoral education is an important issue, and a topic that educators in many countries are working on. While the specifics may be different as a function of the type and format of doctoral education, the principles of assessment will likely be the same.

Having been involved in both graduate education and doctoral education specifically, and within the traditions of graduate education in the United States, my thinking is grounded within our system of education generally. I emphasise this because in establishing and expanding doctoral programmes in nursing, educators have been guided by the established frameworks for this level of study in other disciplines, and the policies and practices that graduate schools have developed on research universities. For it is only in this manner that we in nursing could hope to become an integral part of the fabric of higher education, and be viewed as peers by other disciplines. It is most likely that you have done the same, and have developed your own graduate and doctoral programmes within the traditions of your own system of higher education within established university structures. It is important to bear this in mind in the process of developing assessment structures.

To give you an idea of how rapidly doctoral education has expanded, developed and evolved within the US, consider that when I was in doctoral study there were only five programmes. Now, 28 years later, we have 75 doctoral programmes in nursing. This rapid expansion has serious ramifications for quality. The number of nurses with doctoral preparation to fully meet the demand has not matched the pace at which expansion has occurred.

Educators in each of our societies face challenges. Some of ours are: dwindling financial resources, shifts toward part-time study, changes in the role of nursing, especially in advanced practice roles, rapid technological developments in health care,

increasing public and institutional scrutiny, declining student enrollment, among others. In order to meet these challenges we must rationally identify our strengths and weaknesses, set goals, allocate resources and evaluate the responsiveness of programmes to internal and external interventions and changes over time. Accomplishing these tasks requires being able to measure programme quality. The concerns have become more pressing in light of scarcer resources, the need to attract high quality faculty and students, institutional support, and demands for accountability for educational programming and resource utilisation (Ketefian & Hagerty, 1987). Educators in each country may want to identify their own challenges and how they are responding to them.

Historically, programmes have based their determinations of quality on reputational rankings, resource acquisition, curricular content and more recently, on outcome achievement and student involvement in learning. Each of these approaches makes use of different criteria and provides different perspectives on quality. Further, none of them singly will be sufficient for assessing programme quality and effectiveness.

GOALS OF GRADUATE EDUCATION

A good starting place is to determine the main goals of graduate education. Within the US the Council of Graduate Schools in the United States defines graduate education as: "post-baccalaureate programmes [of a discipline] and research leading to master's and doctoral degrees. The functions of graduate education are the acquisition, transmission, preservation, application and refinement of knowledge" (Graduate Council, Council of Graduate Schools in the United States, 1977). Within this definition, programmes have developed with different areas of emphasis. In professional fields, at the master's level there are typically some elements of research but mostly, the focus tend to be on advanced practice. In nursing, historically we have focused on role preparation, such as teaching, practice, or administration. More recently, most programmes have been exclusively emphasising advanced nursing practice. Programmes with an administration focus have been offering this as an area of specialty rather than as a "role." These changes have come about as a result of recent guidelines clarifying the purpose of master's education in nursing, in view of changes that have occurred in health care (American Association of Colleges of Nursing, 1996).

At the doctoral level, there used to be some "role preparation" as well, such as for education or administration. However, this is no longer the case. It is generally accepted that doctoral programmes should prepare leaders, and their focus should be on theory development for nursing, generation of advanced knowledge for the field, and research training. This change has come about as a result of numerous national discussions, which again have been crystallised in the form of "quality indicators" prepared by the AACN (1993).

It needs to be understood however, that within these general frameworks there is much variation in graduate education across the country, and schools tend to capitalise on the strengths of their faculties. Thus, numerous areas of new specialty and programme foci have emerged in recent years. While these indicators are generally well received, not everyone is satisfied. For example, undergraduate programmes that focus on teaching, that are hiring doctoral graduates, have been expressing dissatisfaction with the exclusive focus of doctoral programmes on research training. These schools would like to see a change so that there is diversity in doctoral preparation, and would like to see a heavier emphasis on the teaching role and faculty functions so that their needs are better addressed.

CRITERIA FOR QUALITY ASSESSMENT

Various authors writing on graduate education have proposed criteria for quality assessment. These generally address faculty, students, programmes, and resources, with some variation. In viewing models of evaluation of quality within health care, the one proposed by Donabedian (1980) related to structure, process and outcome could be a serviceable framework for us as well. Using the framework from health care, we can place the domains of quality along one axis, and the four criteria along the other, creating a matrix. This matrix can then be a good guide in identifying specific quality indicators for each of the four criteria in each of the three domains (structure, process, and outcome).

The "Indicators" of quality (American Association of Colleges of Nursing, 1993), have served to guide us, whether we are continuing programmes or for institutions planning to establish new doctoral programmes. The areas covered are Faculty, Programme of study, Resources, Students, Research, and Evaluation mechanisms.

The section on faculty addresses qualifications, diversity of intellectual perspectives, active participation in research and its utilisation within the broader community, and mentoring of students.

Programme of study deals with core programme elements such as research, theory and knowledge development, history/philosophy of science and research methods, and includes recognition that the programme emphasis is determined by the faculty's area of research; it also addresses cognate study and opportunities for role development.

The resource section deals with sufficiency and variety of resources and services to faculty and students to meet programme goals, and the appropriate access of nursing faculty and students to overall university resources.

The section on students deals with the applicant pool, student selection, their diversity and qualifications, match between faculty/student interests, preparation of students for leadership roles, their training in research and their ability to obtain awards.

The section on research addresses the value placed on research and the need for the institution's mission and reward systems to explicitly recognise it. Research is expected to be programmatic, and there should be institutional and peer supports in place.

The section on an evaluation plan addresses use of a variety of types and sources of data for a comprehensive plan to be implemented. It highlights the need to generate data for programme improvement as well as for comparative purposes; the availability of resources to carry out the evaluation plan is also addressed.

A comment on criteria: I believe it is important that the criteria used be those that are generally accepted by a given credible body, and are publicly known. This may be a government agency, the university, or professional group. Otherwise we may be evaluating something that has no external referent and may be unique to us. However, adherence to an established set of criteria and indicators need not preclude inclusion of indicators that uniquely interest you. For example, our graduate school periodically examines the time it takes for departments to graduate their students from the time of entry. Nursing has never used such an indicator, but when submitting reports to the graduate school, we now regularly provide information on this item. It enables the graduate school to compare time to degree across all departments of the university. It gives them a concrete indicator of measurement, even if we do not look at this indicator for self-assessment purposes.

EXAMINING THE VALIDITY OF AN ASSESSMENT FRAMEWORK

We can now return to the assessment framework in matrix form, described earlier, and review the "indicators." It is possible to fit each of the indicators in specific cells in the matrix. The research indicators fit either with faculty, programme, students or resources, and therefore, research does not appear as a separate criterion. Similarly, the indicators regarding an evaluation plan are part and parcel of other criteria, and are integrated throughout.

Examples of indicators applied to the matrix: In order to see how the indicators might fit in the matrix structure, one can test the idea by using a few examples from each category of indicators related to the criteria previously identified. These examples are inserted in the appropriate cell in the matrix presented in Figure 1.

Faculty credentials and intellectual diversity, fits in cell A, faculty conduct research, cell B, faculty assist in resource generation, cell D, J & K, faculty mentor and socialise students, cell H.

Most **programme of study** elements fit in cell E. Students attain independence as researchers, cell I, academic policies are clear and in writing, cell L. Most **resources** fit in cell J.

Students: Pool is qualified, diverse, and applicants meet admission criteria, fit in cell G. Students obtain awards, cell I. Students have leadership experiences, cell H. Programme completed in a timely way, and students co-author papers with faculty, cell I.

Research is central to mission of the university, fits cell A, and is part of reward system, cell L. Research is programmatic, cell C. Review

Figure 1: Assessment Practical Framework with examples of Quality Indicators

CRITERIA	FOCUS Indicators for each Area of focus		
	Context/Structure/ Input	Process/ Environment	Outcome/ Product
Faculty/Research	**A**—Faculty have appropriate credentials & are intellectually diverse Research is central to university mission.	**B**—Faculty conduct research. Review mechanism in place. Sufficient research in department to justify offering doctoral degree.	**C**—Research is programmatic. Sufficient research in department to justify offering doctoral degree.
Program	**D**—Faculty assist in resource generation.	**E**—Focus on nursing science, research, cognates.	**F**—Evaluation is of high quality, is systematic, & makes use of variety of data.
Students	**G**—Applicant pool is qualified, diverse & meets admission criteria. Admission criteria are consistent with that of parent institution.	**H**—Faculty mentor & socialise students. Students obtain leadership experiences.	**I**—Students attain independence as researchers. Students obtain awards. Program completed in a timely way. Students co-author papers with faculty.
Resources	**J**— Faculty assist in resource generation. Human, financial, technical & library resources are sufficient.	**K**— Faculty assist in resource generation.	**L**—Academic policies clear & in writing. Research is part of the university's reward system.

mechanisms, cell B. There is sufficient research in department to justify offering doctoral degree, fits cell B and C.

Evaluation plan. The elements fit in cell E and F.

Examples from Ongoing Practices

In an effort to illustrate some of these principles at work, I present examples from our own evaluation procedures that take a systems approach. These are only illustrative and it would be impossible to present an exhaustive list of all that we do for quality assessment purposes.

Students
Admission. We have six criteria on which we evaluate applicants: The general test of the Graduate Record Examination, undergraduate and graduate grade point averages, scholarly writing, recommendation letters, and statement of goals. For each of these we have developed a rating scale. Each faculty member of the admission committee evaluates each of the applicants on these criteria, and we summarise these as the basis for later discussion and make overall judgments as a group. The first three measures are quantitative and are easy to rate. For scholarly writing, we look for ability to express thoughts, to think abstractly, independently and creatively, and the quality of presentation. For recommendation letters we look for strength of recommendations, consistency in views of various referees, views on ability to engage in doctoral study, and appropriateness of choice of referees. In the statement of goals we look for consistency with school and university goals and resources, consistency with programme philosophy, degree of clarity in which goals are presented, and evidence of maturity and thoughtfulness in which goals are expressed. A member of the admission committee interviews each applicant. While this is not a distinct criterion, we are seeking additional information through the interview to enhance what we can learn about the applicant from the written record. The interviewer will look for goal directedness for doctoral study, what field of interest applicant has, learn about types of professional experiences the person has had, ability to communicate verbally, what the student believes her/his strengths and limitations to be, and any comments of note by the interviewer.

Progression. Reviews of grade point average and faculty assessment of student performance are ongoing and occur each semester as well.

Graduation. Dissertations are evaluated by a peer group of faculty and the graduate school.

Programme of Study - Curriculum
The School of Nursing has an elected curriculum committee representing every programme, students, alumni, and the nursing community. The committee has criteria, policies and procedures in place approved by the entire faculty. All programmes are reviewed every five years. In addition, all new and revised courses/programmes are reviewed, and then presented to the entire faculty for their approval.

Faculty
Faculty hiring criteria and procedures are in place and are monitored by the university.

Annual review is carried out through self-review, followed by the department chair and the dean. Review criteria address the extent to which, in each year, faculty has contributed to the tri-partite mission of the University, research, teaching and service.

At the time of approval of student dissertation committees, the director of the doctoral programme considers a faculty member's programme of research, and whether the faculty member has shown growth in the performance of research supervision and teaching competence. In order to do this type of assessment it is important for the director to be familiar with faculty publications, to

consider quality of previous dissertations supervised, and how these graduates are performing after graduation (such as with regard to publications, grants, regional or national awards and recognition received).

Resources

Resource evaluation is ongoing. Any needs or department chairs, faculty, students or the programme director can identify problems. Appropriate persons or offices then address the identified problems.

Review by the Graduate School

The Graduate School staff reviews many of the above aspects. The graduate school reviews student qualifications at entry to the programme. They also review dissertations, the curriculum and courses, faculty qualifications, completion rate of students, sufficiency of resources, length of time it takes for students to complete, among other issues.

Goals of Programme Evaluation

One purpose of evaluations is to enable ongoing self-assessment and improvement of what we do. The presentation to this point has focused on this aspect.

Another important purpose of evaluation is to enable comparison with others. Institutions, prospective students, and indeed, the public want to know. This comparative analysis has two aspects: one is comparison with other programmes on campus on common criteria. The second is comparison with other programmes in the discipline. Therefore, in order to make this possible, it is necessary that the criteria and the majority of the indicators used be determined by an external group. This may be a professional group, or the University's Graduate School. It is therefore necessary that we produce evidence of our performance on a recognised and generally agreed-upon set of criteria and indicators. This does not

preclude our developing indicators to evaluate any particular areas of specific or unique interest to us.

Comparison with Other Campus Programmes

Graduate schools in the United States conduct various studies and provide their departments with information. In addition, they ask programmes to compile and provide data on specified criteria and indicators. These are then published for all to see. The funding level and support to departments for student fellowships and financial aid are determined on the basis of such evaluations. Many graduate schools also conduct reviews of PhD programmes by inviting external evaluators. The programme to be reviewed prepares a self-study report; this is then followed by a site visit.

Comparison within the Discipline

We are all using the AACN quality indicators. However, it is not required that we do so. Furthermore, in most cases no quantitative data are generated, and there is no requirement that any assessment we conduct be sent to a location for comparative purposes. Given reviews by university graduate schools, nursing has not created a mechanism such as accreditation to evaluate doctoral programmes. The nursing accrediting bodies review only bachelor and master's programmes. The reputational studies and ratings that the media conduct from time to time have been serving the purpose of informing the public about the relative standing of programmes in various fields of study.

One mechanism is available for programme evaluation with use of an externally determined tool, with the possibility of comparison across institutions in a meaningful way, and that is the Graduate Programme Self-Assessment (GPSA). It is administered by the Educational Testing Service (ETS), which is the same body as that which oversees the Graduate Record Examination, a test generally used by graduate schools. The GPSA is designed to be applicable to any discipline. It

enables the development of up to 20 questions by a school, "local questions," to address any specific areas not covered by the questionnaire. GPSA versions are available for undergraduate programmes, master's programmes and doctoral programmes. Both the doctoral programme and master's programme versions have 16 dimensions, and questions that address many areas of interest in doctoral programme evaluation.

The 16 dimensions of GPSA are: environment for learning; scholarly excellence; quality of teaching; faculty concern for students; curriculum; departmental procedures; available resources; student commitment and motivation; student satisfaction with programme; student assistantship or internship experiences; departmental direction and performance; faculty work environment; alumni dissertation experiences; value of educational experiences for employment; faculty research activities; faculty professional activities.

If more than 10 schools in a discipline use the GPSA, ETS will make comparative data available by ranking the schools on each of the dimensions, and will provide this upon request. The reason for this caution is to make sure the confidentiality of schools is maintained. However, the major concern is that use of the GPSA is voluntary and it is not known how many schools are using it. We do use the GPSA every five years. From time to time, ETS has given us comparative data if we happen to be conducting our review in a year when at least 10 programmes in nursing have used it.

A nice feature of this instrument is that each school can use the data generated from the GPSA to track its progress over time on each of the dimensions. Information on the GPSA is very detailed and needs to be studied with care by anyone wishing to use it. There are some other types of indicators and data that are typically obtained that are not mentioned in the matrix and the criteria by the AACN, but could nicely fit in the practical framework. For example, satisfaction of students fits in cell I (see Figure 1), faculty satisfaction in cell C, satisfaction of employers with the performance of graduates in cell I, student grades and pass rates in cell I, length in programme and completion rates in cell I.

It is possible, indeed necessary, to break down some of the indicators into smaller, measurable indicators. An example of this type is the quality of the dissertation. One would want to evaluate it on contribution to science, creativity in approaching the problem, quality and quantity of the research represented, its scholarly character, among others. Another area is the composition of the dissertation committee. We might look for things such as: whether faculty members are of sufficient rank and experience to provide high quality supervision; whether the members, most importantly the chair, are productive scholars; whether the members possess different areas of expertise required for the particular investigation; whether there are different intellectual points of view represented.

In making overall judgments about quality one must decide about the relative weighting of these criteria and indicators. It is generally agreed that the most crucial factor in quality in doctoral education is the faculty – their credentials, their research, and their mentorship of students.

The framework described for doctoral education may be applicable to undergraduate and master's education. The main differences would lie in the indicators chosen. Also, one must make certain that the indicators are valid and congruent with the goals and objectives for the level of study.

Suggestions for Devising an Evaluation Scheme

Given that in your own system of education there are no formal course requirements, you have the challenging job of developing criteria and indicators that enable you to follow and evaluate student progress. However, as you can see from the general

description of the sample criteria and indicators presented, there is much that can be done.

For institutions that have coursework as part of doctoral study, all the areas covered within the matrix will be relevant for your consideration. For those schools that do not require coursework, all but a few areas would be pertinent. In fact, the only exception that may not be relevant to evaluate are some but not all of the process elements for the programme of study. I am aware that some institutions in Europe, which offer doctoral programmes without coursework, are engaged in similar efforts to develop criteria and procedures for quality assessment (for an example of this, see Hallberg, 1999).

So, why should any of us do this? Given that programme assessment is time-consuming, takes both commitment and resources, and involves every person involved in the educational enterprise, one must have good reasons for engaging in these processes. Here are some of the reasons why it is important to engage in evaluation.

- It enables you to determine if your programme goals are being met.
- It enables you to assess whether society's needs are being met.
- It gives students feedback on how they are performing.
- It gives you ongoing feedback on your performance and enables you to make appropriate corrections and changes. This is now called Continuous Quality Improvement.
- It demonstrates your accountability in justifying use of public resources.
- It informs your profession, your institution, and the public as to what you are doing, to what end, and how successfully you are doing it.
- It enables an informed allocation of resources.
- It assures quality.

There is much room yet in the area of developing more specific outcome and process measures, as well as research regarding these. For example, while we have some idea of what makes for a rich learning environment, we are not sure of the precise elements that account for and contribute to student learning. How can these elements be identified, cultivated, and is there a direct, or indirect, correlation with what types of specific student outcomes?

Another example of where we still need to do more work is in assessment of both quality and impact of faculty and student research. Quality of the research itself is easier to measure as there are specific and measurable criteria; what of significance and impact? What exactly does this mean?

There are still ambiguities and uncertainties to be resolved, and much work yet that needs to be done. However, we cannot wait until all answers are in. Ultimately, what are at issue here are the interests of nursing and society. The questions to be asked are: are the faculty and students contributing to nursing science and to your nation's health through their research? Are they translating nursing research to practice to improve the quality of care your citizens receive?

Doctoral programme evaluation is an area where we might all benefit from international collaboration. I invite us all to work together within the framework of the newly developed group, the International Network for Doctoral Education in Nursing [URL: http://www.umich.edu/~inden/].

REFERENCES

American Association of Colleges of Nursing 1993: Indicators of quality in doctoral programmes in nursing. Washington, DC: Author.

American Association of Colleges of Nursing 1996: The essentials of master's education for advanced practice nursing. Washington, DC: Author.

Donabedian, A 1980: Explorations in quality assessment and monitoring, Volume I: The definition of quality and approaches to its assessment. Ann Arbor, MI: Health Administration Press.

Graduate Council, Council of Graduate Schools in the United States 1977: The doctor of philosophy degree. Washington, DC: Author.

Hallberg, IR 1999: Quality development in the nurse's doctoral education from a Swedish perspective. Paper presented at the International Network for Doctoral Education in Nursing, London. (Paper can be viewed at Website www.umich.edu/~inden/)

Ketefian, S & Hagerty, BK 1987: Putting master's programmes to the test. **Nursing and Health Care**, 8(9):503-507.