Evaluation of an Online Curriculum:
Defining Standards of Quality in Graduate Nursing Education

Melissa D. Avery, CNM, Ph.D.
Associate Professor
School of Nursing
University of Minnesota
5-140 Weaver Densford Hall
308 Harvard St. SE
Minneapolis, MN 55455

Bradley A. Cohen, Ph.D.
Assistant Director and
Senior Educational Technology Consultant
Digital Media Center
212 Walter Library
University of Minnesota
Minneapolis, MN 55455
(612) 626-0282

J.D. Walker, Ph.D.
Coordinator of Research and Evaluation Services
Digital Media Center
212 Walter Library
University of Minnesota
Minneapolis, MN 55455
(612) 624-1097

Acknowledgement

This project is supported by funds from the Division of Nursing (DN),
Bureau of Health Professions (BHP), Health Resources and Services
Administration (HRSA), Department of Health and Human Services (DHHS) under grant
numbers D09HP00115, NM, WHCNP, and PHN Graduate Education Via Technology,
for $1,630,100 and D09HP04068 Technology Enhanced Learning in Graduate Nursing
$966,600. The information or content and conclusions are those of the authors and
should not be construed as the official position or policy of, nor should be any
endorsements be inferred by the Division of Nursing, BHP, DHHS or the U.S.
Government
Recent years have witnessed a proliferation of online course offerings at a wide range of institutions of higher education. Many graduate programs are now offered entirely online, particularly in professional schools. Nursing education is no exception. A number of authors have described the use of online strategies in graduate nursing programs (Mills & Reese, 2000; Mills & Hrubetz, 2001; Jesse, Taleff, Payne, Cox, and Steele, 2006). One notable trend in this growth of online education is that online course offerings have outpaced the evaluation of the quality of those offerings, particularly at the program level.

In this paper we discuss an evaluation project undertaken by nursing faculty at a large Midwestern U.S. university and educational technology professionals from the same institution. This interdisciplinary group collaborated in the development of standards of quality for online graduate nursing courses and the evaluation of 16 online graduate nursing courses using those standards.

Background

In 2000, nursing faculty received federal funding from the US Department of Health and Human Services (Health Services and Resources Administration, Bureau of Health Professions, Division of Nursing) to develop online graduate nursing courses in three specialty areas: nurse-midwifery, women's health nurse practitioner, and public
health nursing (Avery, Ringdahl, Juve and Plumbo, 2003). Additional funding has provided for the re-design of two more graduate specialties, psychiatric mental health clinical nurse specialist and nursing and health care systems management. Educational technology professionals from the university’s Digital Media Center met with graduate nursing faculty to plan for the evaluation component of the grant.

The evaluation process had a number of goals beyond simply satisfying a reporting obligation including:

- discovering areas of potential improvement both for individual courses and across multiple courses;
- identifying examples of excellent online teaching within the School of Nursing to provide guidance for faculty seeking to develop new or improve existing online courses;
- developing an evaluation instrument that could be understood by faculty members and instructional designers, and used to evaluate online courses in a reasonable amount of time;
- providing a checklist of the important components of quality online education for faculty in the process of developing or revising online courses; and
- initiating important dialogue within the School of Nursing regarding what faculty consider to be the most important elements of good online courses.
We began the process with the goal of determining baselines or benchmarks for judging success. Those benchmarks would be grounded in a set of shared beliefs concerning effective course design and quality instruction, influenced by a number of factors, including the culture of teaching in the school; the disciplinary content of nursing; and characteristics of nursing students taking the courses (their learning styles, backgrounds, expectations, etc.). The articulation of this set of shared pedagogical beliefs followed a survey of relevant literature related to online pedagogy (Cobb, 2001; Billings, 2001; Phipps, 2000; Wright n.d.), and an examination of existing instruments for the evaluation of online courses. We also drew on the expertise of faculty members in the school who had the most experience with online teaching. The result was the explication of several standards that were well-grounded in existing literature and that fit with the beliefs of the faculty involved. These standards were organized in several categories:

- **course mechanics**: the clear articulation of course goals and objectives, technical requirements, pre-requisites, time commitment, etc.; the connection of course objectives with the learning activities and assessment mechanisms of the course.
- **course organization**: ease of navigation; scaffolding of course material and activities from simple to complex; accommodation of multiple learning styles.
- **student support**: availability of faculty and/or TAs to address student needs; support for the learning activities of the course; access to technical support.
- **communication and interaction**: degree and robustness of student interaction with other students, faculty members, and course material.
Methods

Evaluation instrument

The existing literature was reviewed to locate an instrument designed to measure these standards that could be used both retrospectively, to evaluate online courses that had already been taught, and prospectively, by faculty engaged in designing, developing, or revising an online course. Although ideas were drawn from existing evaluation instruments (including the Robert Wood Johnson Foundation’s Partnership for Training project, Quality Standards for Learner-centered Online Instruction, no longer available online), none were able to be used for our project because they were written in instructional design jargon, making application difficult for faculty members; they were extremely long and unwieldy, requiring extensive time to apply to a single online course; or did not seem to reflect the school’s standards for high-quality online instruction.

Through a series of meetings, the team of faculty and educational technology professionals constructed an evaluation tool (online at http://dmc.umn.edu/nursing-evaluation/) which reflected the standards described above and required approximately two to four hours to apply to a single online course. The tool initially consisted of 20 items to be rated on a 5-item scale (higher number represented closer adherence to the standard) and one general open-ended item. Raters were encouraged to add comments to explain their numerical ratings. The scale items were formatted as in the illustration below.
13. Course content is presented to appeal to a variety of learning styles.

- Instructor uses varied activities such as case studies, virtual lab, interactive simulation, cooperative projects, self-tests.

<table>
<thead>
<tr>
<th></th>
<th>To a very small extent</th>
<th>To a small extent</th>
<th>To a moderate extent</th>
<th>To a great extent</th>
<th>To a very great extent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comments:

We evaluated the reliability and validity of our instrument by means of a pilot test that focused on a graduate ethics course that had been taught several times by single faculty member. Three different faculty members and two educational technology professionals used the instrument to rate the pilot course by examining an archived version of the course’s WebCT site. Student assignments and discussion postings within a threaded discussion tool, as well as the course materials were available for review.

The raters met with the instructor to review the ratings, identify aspects of the course that may have been missed by the raters, and generally debrief about the course rating experience. Calculations of inter-rater reliability were not computed, but the raters discussed how to “calibrate” the instrument by establishing agreement on the meaning of the 20 items. Faculty and educational technology professionals had divergent ratings on a
few items; this was resolved by discussion resulting in a common understanding of what each question was asking and agreed to by all five raters. Face and content validity were also determined through this pilot process. Instrument use and review by faculty experienced in both classroom and online teaching, as well educational technology/instructional design professionals, provided the background and expertise to assess this instrument.

Following the pilot process, minor clarifying changes were made to the instrument and the evaluation process as a whole was modified. In addition to activity on the course web site, students communicated with the professors by email or phone, and typically had one or two face-to-face meetings. Substantial interaction occurred among students and between students and instructors during these face-to-face meetings and was not available for review. Therefore, we altered our process by adding a follow-up interview with the instructor of each course evaluated.

_Evaluation procedures_

The revised tool was applied to each of the 16 courses. These included courses specific to public health nursing, nurse-midwifery and women’s health nurse practitioner specialties as well as core courses required for most students in the master’s program. Two nursing faculty members and one educational technology professional examined each course. One faculty member was part of the funded project and had worked on the tool development; the others were drawn from a larger pool of experienced online teachers. Data from all raters were compiled and reviewed, shared with each respective course instructor(s), and an interview was conducted with each instructor.
In addition to standard quantitative analyses of the numerical data, we independently reviewed the qualitative data (reviewers’ comments entered following each question on the instrument, and notes taken during the follow-up interviews with the instructors) looking for best practices, common themes, points of clarification, and areas of confusion and disagreement among reviewers. Each author independently reviewed the data and described the primary information related to each question. Then common themes that emerged across all the questions were identified. The three authors met and discussed their individual interpretations, and came to agreement on identified themes. A formal coding and categorization scheme was not utilized.

Results

*Highest and lowest mean scores*

Items on the instrument that yielded the highest and the lowest mean scores across all 16 courses were reviewed and are shown in tables 1 and 2. The overall mean score for all items was 3.91 on a 1-5 scale.

Table 1

*Highest Mean Scores*

<table>
<thead>
<tr>
<th>Question (Q)</th>
<th>Mean score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Q2: Goals and objectives appropriate to level of the course 4.51
Q1: Goals and objectives clearly stated 4.40
Q9: Learning activities are clearly described 4.40
Q10: Learning activities applicable to the course are sufficiently supported 4.34
Q12: Evaluation mechanisms/instruments measure objectives 4.28
Q4a: Learning activities and teaching strategies match the course objectives 4.26

Table 2

*Lowest Mean Scores*

<table>
<thead>
<tr>
<th>Question (Q)</th>
<th>Mean score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q4b: There is a written connection between the course objectives and learning activities</td>
<td>2.88</td>
</tr>
<tr>
<td>Q7: Technical requirements specified with regard to both skill and equipment</td>
<td>3.16</td>
</tr>
<tr>
<td>Q8: Realistic time commitment related to credit load</td>
<td>3.26</td>
</tr>
<tr>
<td>Q3: Pre-requisite or prior knowledge required for course outlined</td>
<td>3.38</td>
</tr>
</tbody>
</table>
Inter-rater reliability

As a check on the consistency of our evaluation procedures, we examined the variability of ratings given to all courses by the different raters. Educational technology professionals gave higher mean ratings than nursing faculty reviewers on 16 of 21 items across all courses. These differences were statistically significant (at the p < .05 level) in four cases:

- Question 2: Goals and objectives appropriate to level of the course
- Question 5: Faculty availability to address student needs
- Question 12: Evaluation mechanisms/instruments measure objectives
- Question 15: Flexibility for student input to shape the course as appropriate

Several factors may explain the discrepancies. Nursing faculty may be more critical of courses in their own area than outsiders by virtue of their greater subject knowledge and/or their experience teaching similar courses. Educational technology professionals may have been less critical as outsiders and not wanting to presume or offend. Finally, the educational technology professionals have reviewed more online courses than the faculty members, and the widely varying quality of these courses may have provided a comparison against which the nursing courses were impressive.

Themes from qualitative data

Match between course objectives and learning activities. One question asked specifically about the connection between learning activities present in a course and that
course’s stated objectives. In other words, are the things students do in a course well-suited to help them achieve the course’s learning objectives?

Many of the courses we examined included learning objectives that were relatively high in Bloom’s taxonomy (Bloom, 1956). Instructors hoped that students would not only acquire knowledge and comprehension, but that they would be able to apply that knowledge, and engage in analysis and synthesis. In many cases it was not clear to raters how the learning activities, which gave students little practice at application, analysis, or synthesis, would help them achieve those higher-level objectives.

A nursing theory course with an assignment where students were introduced to criteria used to analyze conceptual models/theories and then asked to complete an analysis performed especially well. The assignment matched a corresponding high level course objective and in addition, the professor described to students why this skill was important and how it linked to clinical practice. In contrast, a clinically focused course had a number of lower level objectives in each content module, but it was not clear how those objectives and corresponding module activities related to the overall higher level course objectives. Determination of achievement of these objectives may be made in the clinical setting, but this was not completely clear as described in course materials.

Technical competence. Students are increasingly sophisticated users of technology and have little patience for courses that are cumbersome to navigate and not easy to use (Jones, 2002; Ernst et al, 2005). Online courses must therefore be easy to use,
consistent in the information provided, free of errors as much as possible, visually appealing, and links to outside information must work. A common mistake that makes navigation more difficult is insufficient use of hyperlinks within the course, specifically describing instructions for a learning activity located in another part of the course and not providing a direct link.

Errors or inconsistent information were found in a few courses we reviewed. Navigating around a course was not always easy or intuitive. However, some courses were complimented as providing a structure that was very easy to use. Technical requirements of students in advance of taking an online course must be clear. This information —whether and how it was provided—varied considerably across all courses. One good example was information that students would need access to and skill in using Power Point to complete a course assignment including a link to an outside tutorial for students who may not have the skill.

**Clarity of support.** When students are provided or referred for support in order to achieve the objectives of the course, the information must be provided in an easily accessible and understandable format. For example, if students need to watch a video or listen to an audio file, they must be provided information about needed plug-ins or access to other information. If students are required to create and post online presentations, clear information is needed so they can do the work productively and in a timely fashion.
Comments related to this theme included the importance of clear information about how and when the instructor is available for student support. Some instructors provided this information prior to the start of the course via email or face to face at the start of the course. Others included support information within the course site in a specific location and/or along with specific learning activities and assignments.

One really good example was a course that provided a link to an outside resource called “Is Online Right for You?” This site provided students with an assessment of their readiness and disposition toward online academic work, information to assist novices understand what to expect, and tips on getting started in an online course. In contrast, some courses expected students to complete a PowerPoint presentation; but no statement was made about the requirement for those skills, nor was support provided.

*Learning styles.* Another salient theme that emerged was related to learning styles. Online nursing courses used a great variety of different learning activities (such as group work of various sorts, scavenger hunts, self-tests, letter-writing, and talking circles), but they tended to be text-heavy and to make little use of varied media.

Examples of good use of media include the use of a photograph and audio recorded message from the faculty member delivering a welcome message to students, graphic images inserted into text-based content making it more visually appealing, and audio with PowerPoint presentations so students can listen to the faculty member present information on a particular topic with visuals.
**Student voice in the course.** Two important components of graduate education are student input into the course and higher level learning outcomes such as analysis and synthesis. Some clinical courses will be heavier in specific detailed information related to advanced practice specialty content and students will have less natural opportunity to contribute to the conduct of the course. However, some courses did or could have offered students a chance to pose questions, provide interpretation or synthesis of content, or suggest new areas of discussion. As with the theme of learning styles, student voice will be more apparent in courses with more varied learning activities.

Several good examples of student voice included asking student to reflect on ways they observe research being integrated into their current practice settings and opportunities for further use of research in practice, asking students to bring two questions for discussion at a face-to-face meeting, and asking students to complete a brief mid-term evaluation to help shape some aspects of the remainder of the course.

**Interaction.** The degree of student-student interaction is an important component of online education. Several recent meta-analyses of studies of distance education courses conclude that the presence of student-student interaction in an online course is a strong predictor of learning outcomes in that course (Bernard, 2004; Hiltz, 2000; Zhao 2005).

While all online nursing courses used online communication tools (email, threaded discussion, synchronous chat), the degree of genuine interaction among students
was often quite limited. True interaction occurs when one person’s actions are conditioned by the actions of another person (Wagner, 1994; Wagner, 1997; Thurmond 2004), and this was not observed in a number of the courses. In online asynchronous discussions, students would often just say their piece as required by the assignment and leave, without taking account of or reacting thoughtfully to anyone else’s contributions. One good example of students effectively engaging in conversation was a research assignment where students posted a summary of several steps of an assignment they were working on and each gave and received peer comments in a small group via the asynchronous discussion tool. Specific instructions provided by the faculty member for this learning activity likely resulted in the increased level of interaction.

Discussion

Guided in part by Michael Patton’s utilization-focused approach to evaluation (Patton, 1996), our data analysis centered on three action categories, representing ways in which we hoped that our findings would be used:

- course-specific items for potential revision by individual faculty (things professors may want to change or work on in their courses);
- issues for School of Nursing faculty to address as a group (items having to do with overall standards of quality for online nursing courses)
- items that reflect a need to improve the evaluation instrument (ambiguities, confusing questions, etc.)
In order to promote the use of our findings, individual faculty were given the data from their course evaluation to use as they desired in continued revisions to their online courses. These data have been presented to the faculty as a whole so that the information can be used school-wide to continue to improve the quality of online courses.

**Highest and lowest mean scores**

A conspicuous feature of the items that received the highest ratings (Table 1) is that several have to do with the goals and objectives of a course, particularly their presentation and their connection with the learning activities and evaluation mechanisms of the course. The school has emphasized the explicit statement of course goals and objectives and their integration into the course in its curriculum development process for a number of years, and this emphasis was reflected in its online courses.

The lowest-rated items (Table 2) all have to do with the provision of certain information to students in a course. Our evaluation revealed many cases in which course Web sites used boilerplate information rather than information tailored to a specific course, and this fact may account for the lower ratings given to three of these four items.

For instance, regarding Q7 (technical requirements specified with regard to both skill and equipment), courses frequently simply linked to standard school or university technical requirements related to students having a computer with certain specifications and level of internet connection. However, students were not always informed about what
they needed for this particular course, such as particular browser plug-ins or multimedia software. Regarding Q3 (pre-requisite or prior knowledge required), courses often simply noted that students must be "graduate students in nursing", rather than describing the background knowledge and skills students need to succeed in a particular course.

Online courses with more specific information would be more helpful to students, but this is a policy issue best decided by the department or school. Boilerplate information can provide a baseline of information for students that holds across courses, but may be misleading in courses that demand more specialized skill and equipment.

Finally, with respect to the lowest-rated items, the scores were relatively positive. Q4 (written connection between course objectives and learning activities) is the only item in the group to score below 3 on a 5-point scale. This could be an artifact of sympathetic reviewers. Q4 is the one item for which there was no boilerplate, and the one item for which the lowest possible scoring was therefore likely.

Making the connection between course activities and learning outcomes explicit facilitates metacognition, in other words, the ability for students to monitor their own learning, (Bransford, Brown & Cocking., 2000) and may improve motivation (Svinicki, 2004). Making these connections can lead to a better match between objectives and learning activities. Policy-level decisions recommending or requiring explicit connections between objectives and activities may facilitate student learning.
Changes to evaluation tool and procedures

The discovery that educational technology professionals’ ratings were systematically different from those given by faculty clearly showed the need to improve evaluation procedures to improve inter-rater reliability. The tool has been revised in an effort to reduce ambiguity. After a review of relevant literature (Roblyer, 2000; Rubric for Online Instruction, 2003; Taggart, 1999), we created a scoring rubric (online at http://dmc.umn.edu/nursing-evaluation/#r) designed to increase the rigor and to reduce the subjectivity inherent in the evaluation process. The rubric provides an abstract description of different levels of success in meeting the criteria for each item in the evaluation tool. These levels were assigned numerical values and accompanied by specific examples, as in the illustration below.

Figure 2. Sample evaluation tool question with scoring rubric.
8. **Learning activities are clearly described.**

- Instructor provides a framework for completing assignments (student discussion, participation, and projects) including expectations for satisfactory completion.

1. The instructor provides very few or no instructions for the successful completion of course assignments, and/or the instructions are of such poor quality or so general and vague as to be of little use to students.

3. The instructor provides at least some instructions for the successful completion of course assignments, but the instructions are uneven in quality, at least sometimes overly general or vague, and/or do not pertain to all assignments.

5. The instructor provides high-quality, specific instructions for the successful completion of all course assignments, such as detailed rubrics for written assignments, samples of high and low quality work, guidelines for working in student groups, study guides for tests and quizzes, instructions for participating in online discussions (quality and quantity of contributions expected), etc.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>To a very small extent</td>
<td>To a small extent</td>
<td>To a moderate extent</td>
<td>To a great extent</td>
<td>To a very great extent</td>
</tr>
</tbody>
</table>
Themes from qualitative data

Match between course objectives and learning activities. The finding that objectives did not always match well with activities in a course could be useful for individual faculty members who want their students to achieve higher-level learning objectives if it is combined with faculty development programs that focus on the use of more sophisticated active learning strategies (McConnell, 2003).

Technical competence. There are two issues pertaining to this theme. One dimension relates to features of the course web site—how intuitive is the navigation and how functional is the site; and the other concerns the degree to which students are prepared with the technology skills they need to succeed in specific courses. We recommend the use of policy-level decision regarding the standard for communicating to students regarding technical competence required for any specific course.

Following this evaluation, the School created of an online orientation site all students must complete before beginning their first online graduate course. The site ensures a baseline level of technology skill among students, and serves to orient them to a template that is used in the design of individual courses throughout the School.

Clarity of Support. Instructors provided information regarding student support resources in a variety of ways. As programs are increasingly available online, this information would ideally be provided in a consistent fashion throughout a program,
across courses. Whatever approach is chosen, consistency and clarity are important to facilitate student learning.

**Learning styles.** Using both a variety of activities and a variety of media will help instructors seeking to provide diversity in their courses, make the courses more interesting and hopefully provide all students an opportunity to interact with the course in a way that is satisfying based on their personal preferences for learning. We have disseminated a number of examples of nursing courses that make excellent use both of varied activities and of varied media, to serve as models for instructors seeking more variety in either aspect of their courses. The evaluation instrument was also revised to clarify this criterion by separating questions about learning activities and use of media.

**Interaction.** Student-student interaction was inconsistent across the courses reviewed. Insufficient guidance for students may have affected lack of robust interaction in some cases, thus overlapping to some degree with clarity of learner expectations. Research about online interaction has consistently emphasized the importance of providing students with explicit structures to guide their interaction (Horton, 2000; Feenberg, 2000; Hiltz, 2000), and while some courses provided students with excellent guidance as to how to use online communication tools interactively, others did not.

Another important consideration with graduate nursing students is their more full personal lives compared with undergraduate students. Our students tend to be older, employed at least part time as nurses, and often have families including young children
that compete for time and attention, reducing the opportunity to converse regularly and in meaningful ways using asynchronous discussion methods. Providing specific instructions and time frames for meaningful interaction and other means for dialogue, including live discussion, should help create opportunities for scholarly dialogue.

We have responded to this finding by highlighting, in written reports and oral presentations, courses that represent demonstrated best practices with respect to creating robust student-student interaction and providing scaffolding (starting with simple activities and building to those that are more complex) for students who are expected to interact in the online environment.

Conclusion

Online teaching and learning is an approach to education that is growing in response to the need to reach students wherever they are located and whenever they are able to participate in learning. As academic units attempt to increase access by moving programs online, there is a need to ensure the quality of those programs. The project described in this paper represents a collaborative development process that produced a set of theoretically informed yet practically grounded pedagogical standards as well as an evaluation tool which operationalized those standards.

This project should provide academic units concerned with quality in online education with a valuable model for quality assurance that examines course mechanics as well as the quality of engagement and interaction in the “virtual classroom”. Our revised
course evaluation tool, with its accompanying rubric, can be of value in the development or redesign of courses for online delivery. Another university department is using the tool for the redesign of all courses in one of its master’s degree programs. The usefulness of the tool in creating and evaluating new online courses will be evaluated. The tool also has the potential to be used to support peer review, either for an entire course or for specific content modules. Similar to a faculty peer attending a face-to-face class session to review the teaching of a colleague, the tool could be used to review a specific content area, providing specific ratings and qualitative feedback.

The evaluation tool and process discussed in this paper is one of a few notable recent examples of program-level evaluation strategies for online programs, including the Quality Matters program (http://www.qualitymatters.org/) the Rubric for Online Instruction project at CSU Chico (http://www.csuchico.edu/celt/roi/), and the University of the Sciences in Philadelphia learning-centered teaching project, (http://www.usip.edu/learningmodules/ltc.htm). Perhaps the fact that recent doctoral dissertations have made evaluation of online nursing programs their focus (Brown, 2005; Brownrigg, 2005; Kaiser, 2005) provides some evidence that evaluation of online learning environments is maturing. Nonetheless, additional research related to quality in the online learning environment is needed to examine student outcomes over entire programs of study, outcomes which should include measures of academic success, professional socialization, and satisfaction with the online learning experience.
References


