Preparing Faculty to Use the Quality Matters Model for Course Improvement

Carol Roehrs
Associate Professor
School of Nursing
University of Northern Colorado
Greeley, CO 80631 USA
carol.roehrs@unco.edu

Li Wang
Training and Development Manager
Amgen, Inc.
Thousand Oaks, CA 91330 USA
lwang01@amgen.com

David Kendrick
Director
Center for Excellence in Teaching and Learning
University of Northern Colorado
Greeley, CO 80631 USA
david.kendrick@unco.edu

Abstract

The number of fully online and hybrid (blended) courses in higher education has increased rapidly in recent years. One factor shown to influence effective online learning is the instructional design of such courses. Because continuous improvement in support of student learning is an important part of online education, many colleges and universities have adopted the Quality Matters (QM) program. QM is based on peer review of courses by faculty members who are trained and certified to assess the design of online courses. They provide feedback to instructors in the form of scores on a rubric and recommendations for change. Another approach to implementing QM standards might be to educate interested faculty members in the use of the rubric so they can review and improve their own courses. This research report summarizes a mixed-methods descriptive study focused on the experiences of faculty participants with different kinds of QM training, self-evaluation of a course, and updating of the course. Qualitative and quantitative data converge to support several main findings about using the QM rubric, identifying and making needed changes without help, wanting help from instructional designers with aspects of course improvement, and needing time in faculty workload to review and improve courses.

Keywords: instructional design, faculty development, faculty training, self-evaluation of online courses, peer review of online courses, Quality Matters

Introduction

Comparison studies have long supported the view that students at a distance learn as well or better than those in the classroom (Western Interstate Commission for Higher Education Cooperative for Educational Technologies [WCET], 2010). The results of these studies are so consistent that Russell (2001) described it as the "no significant difference" phenomenon and provided a compendium of such studies that is still growing (WCET, 2010). One factor shown to influence effective online learning is the instructional design of such courses (Gunawardena et al., 2006; Wiesenberg & Stacey, 2005). The number of fully online and hybrid (blended) courses has increased rapidly in higher education. Because continuous improvement in support of student learning is an important part of online education, many

colleges and universities have adopted the Quality Matters (QM) program. This is a system of faculty peer review using a rubric to assess the instructional design of an online course. The reviews provide scores that reflect the extent to which a course meets the 40 standards and recommendations that instructors use to update course design. Courses can also be submitted to QM to achieve course certification, which is an indication of high-quality design supporting student learning.

In some cases, instructional designers become certified peer reviewers as well as holding an assistive role in the development of new online courses. Even this combination of faculty members and staff may not be sufficient to meet the needs of a university for QM course review. It has been observed that some faculty members who had been teaching online for a time were spontaneously identifying and incorporating QM strategies into existing courses before a review was done. In addition, a number of the QM standards are foundational education practices such as writing measurable learning objectives and linking the objectives to readings, assignments, and assessments of learning. Many faculty members are familiar with those practices. If faculty members understand some QM standards before review or training, perhaps another approach to spreading QM quality to online classes might be to educate interested faculty members in the use of the rubric so they could review and improve their own courses.

This study was based on the premise that faculty members can learn about the rubric, use it accurately to evaluate a course, and then update the course based on their findings to more closely meet the QM standards. The broad research problem was how to improve the quality of online courses offered by the university. While there were many aspects of that problem that could and should be studied, the design of this project was focused on the premise just mentioned. The following overall research question was identified: What are the experiences of and outcomes achieved by faculty members as they learn about the QM model through different kinds of training, and employ the rubric for self-evaluation and updating of an online course?

Background

Quality Matters

Quality Matters (QM) originated from a 2003 grant from the Fund for Improvement of Postsecondary Education (FIPSE) and is a faculty-centered peer-review process designed to enhance the quality of online courses. Now a proprietary system, QM has generated widespread interest. Over 700 institutions have joined QM either as an individual subscriber or as part of a statewide consortium (see QM, 2013 for a list of subscribers). The organization has trained more than 23,000 faculty and instructional design staff (MarylandOnline, 2013a). The essence of QM is peer review of courses by faculty members who have been trained and certified by QM in the use of a research-based rubric that generates numerical scores and feedback comments for instructors. The 2008-2010 version was used in this study; the updated 2011-2013 version is available on the QM website (MarylandOnline, 2013b). The rubric consists of forty specific standards about aspects of course design distributed across eight general standards listed below in Table 1 (MarylandOnline, 2008). Each has a point value ranging from one to three. There are 17 "Essential" standards worth three points each, 11 "Very Important" standards worth two points, and 12 "Important" standards worth one point. To achieve certification, a course must earn a minimum of 72 points including all of the Essential standards. In addition, there are core standards in the rubric to highlight the concept of alignment, which means course objectives should drive the development of learning and assessment activities, and selection of course materials and course technology.

There are two approaches to using the QM program. Informal reviews can be done by one or more peer reviewers from the college or university. Formal or official QM review is carried out by a team of three reviewers that includes a content expert, at least one reviewer external to the instructor's school, and a reviewer serving as the team leader. Formal QM reviews are performed to determine if a course has met the criteria for certification, which provides recognition for the quality of an online course for a fee (then $150 USD). To provide consistency in terminology for the two kinds of reviews included in this study, "peer review" will refer to informal course reviews performed by a single certified peer reviewer from the university, and "QM review" will refer to formal course review by a QM team as described above.

Literature Supporting Peer Review

The employment of faculty peer review in education has long been valued. Cohen and McKeachie (1980) and Millis (2006) are examples of the many authors who have extolled the use of peer observation as a strategy to provide feedback to instructors so they can improve their teaching. Morehead and Shedd (1997) encouraged the incorporation of external reviewers. As online courses have proliferated, so has

attention to the quality and improvement of those courses. Nelson and Van Leeuwen (2008) and Gaytan and McEwen (2007) all supported peer review of online courses while Cobb, Billings, Mays, and Canty-Mitchell (2001) looked specifically at nursing courses. McNaught (2001) devised a systematic approach for peer review of online courses, as did Ross, Batzer, and Bennington (2002). This sort of effort continues with application to whole universities (Adams, Rust, & Brinthaupt, 2011) and specific departments such as nursing (Little, 2009).

Table 1. *QM General Standards (2008 version)*

Category	Title	Content	Points
General Standard 1	*Course Overview and Introduction*	Offers ideas for conducting the initial course overview and introduction to welcome students in an online or hybrid environment. Courses following such suggestions should provide easy and consistent navigation in the website, and clear explanation to students about course structure and content.	11
General Standard 2	*Learning Objectives*	One of the core standards in the rubric that includes some alignment related to course learning objectives. It is suggested that objectives should be written from the students' viewpoint so that they know clearly what they will be able to achieve in measurable and observable terms.	14
General Standard 3	*Assessment and Measurement*	A core standard that addresses assessment of student learning and is related to learning objectives. It is suggested that all assessments should be aligned with course objectives and provide multiple opportunities for students to assess their own learning and gather feedback.	13
General Standard 4	*Resources and Materials*	Related to course materials and resources to support course learning objectives. It is also one of the core standards addressing alignment. It is recommended that all materials should help students make meaningful connections with the goals they are able to achieve.	9
General Standard 5	*Learner Engagement*	One of the core standards, which provides recommendations on the strategies of engaging students in an online and hybrid environment. The design and development of learning activities should support course and module level objectives. Activities help build the online learning community and help students become active learners.	10
General Standard 6	*Course Technology*	Also a core standard that addresses the technology used in the course. Course technology should be current and support course objectives. Instructors need to carefully select course technology that helps students become active learners and provide instructions and access to the technology used.	14
General Standard 7	*Learner Support*	Provides examples of academic and student services resources and support that students might need in the course including the library information, technical support for the Learning Management System, writing center, and other student services information including counseling unit, career center, and so on.	6
General Standard 8	*Accessibility*	Supports accessibility by incorporating ADA standards into the course. Instructors should not only provide ADA policies and describe the supports the campus has available to students, but also demonstrate practices in course design that make course materials available for learners having various needs (e.g., webpage design for impaired vision and hearing).	8

Research Supporting Quality Matters

As the QM process was being developed and tested, Legon (2006) compared the QM rubric to existing accreditation standards, finding good correspondence between them. Recent studies report testing the use of the QM rubric with positive results (Little, 2009; Puzziferro & Shelton, 2009). An update of the research base for the current rubric has been posted (MarylandOnline, 2013b). The literature strongly supports the basic premises of the QM program: research-based standards presented in rubric form, course review by knowledgeable peers, inclusion of external reviewers, a systematic process for course assessment, and provision of feedback to the instructors whose courses are being reviewed.

Method

This was a mixed-methods descriptive study about the experiences of faculty participants with training in the use of the QM rubric, reviewing an online course, and updating the course. This design was chosen because two main categories of data were needed to answer the research question: qualitative data in the form of comments from participants about the processes of training, reviewing, and updating, and quantitative data about the outcomes of faculty efforts in the form of three sets of rubric results for each course. Both sets of data were collected concurrently then used to develop the interpretation of the results in what Creswell and Plano Clark (2010) describe as a concurrent parallel design. A mixed-methods approach accommodates this variety of data and analysis, and provides a more complete explanation of a phenomenon than a single source of data.

Purposive selection of participants was employed to find faculty members who had at least two years of experience teaching online and no experience with QM training. Each was to have a course available that had been offered online at least twice and was not designed using QM principles. After providing consent to participate, faculty volunteers completed a short informational survey. Assignment to one of the three kinds of QM training was based on participants' preferences in combination with available places to create balanced groups.

In order to explore several approaches to faculty training about the use of the QM rubric, three kinds of training were provided: self, short, and long. All participants received the same materials: QM overview and rubric, information on standards alignment, guidance for providing feedback, and the QM Master Review Chart. Self-training participants read through the materials on their own. Short and long training groups shared a three-hour face-to-face meeting to learn about QM, practice assessing selected standards from the rubric, and complete an exercise about the alignment of learning activities and assessments with course and unit objectives. Long training built upon short training by adding three more hours of face-to-face practice. Both of those training sessions were taught by a QM staff expert who was not otherwise involved in the study. None of the groups was offered asynchronous discussion as part of their training to facilitate learning or answer questions since the focus of the study was independent use of the QM rubric after training. The expectation of the participants was to listen and learn during the training, participate in the practice exercises, and ask questions until each felt comfortable using the QM rubric. There was no particular assessment employed to evaluate the level of competency that may have been achieved during the training sessions, nor was previous knowledge of strategies that appear in the rubric assessed.

After training was completed, participants evaluated their own courses using the QM rubric then received peer-review feedback. The participants then made improvements in their courses without help from instructional designers. The revised courses were submitted for official QM review. After the study ended, additional assistance was made available to participants who wanted to continue to improve their courses to attain certification.

Comments about the faculty participants' experiences were solicited at each step of the study to provide qualitative data; recommendations from course reviews were also part of the qualitative data, but content analysis of the recommendations was not completed when it became apparent that the structure of the rubric made the comments very similar for each standard. The scores from rubrics prepared by participants, peer reviewers, and QM review teams provided numerical data for analysis with descriptive statistics; the small sample prevented more powerful statistical analysis. The informational survey provided demographic data for description.

Results

Initial Survey: Participant and Course Data

Six faculty volunteers made up the final sample, with two each in the self, short, and long training groups. Four of the six were from Nursing and one each were from Art and Theatre. There was one male and five females; ages were evenly divided between 30s and 50/60s with a mean of 43. Years teaching online ranged from one to 10 years with only one above five years; the mean was 5.4 years and 3.4 years if the most experienced instructor was omitted.

The courses that were reviewed included one junior level and three senior level courses from Nursing and two at the freshman level from Art and Theatre. Courses were chosen by participants for review because they were part of the instructors' assignments and/or needed to be updated. Five of the courses were three-credit, fully online courses that correspond to didactic courses on-campus. The sixth was a hybrid course representing one component of a 10-credit clinical nursing course.

Looking at the Experiences of Participants

The faculty participants were asked to reflect on their experiences and share their comments several times: after their QM training, after self-evaluation of their courses, after receiving peer-review feedback, after updating their courses, and at the end of the study after receiving the official QM review. Rather than holding face-to-face focus groups, which were perceived as onerous, online discussion forums were set up so participants could asynchronously develop discussions about their experiences with others who had the same kind of training; several questions were posed in each forum. In reality, the participants were not able to comment at these exact points in time, and most caught up after completing one or more of the next steps in the process.

- *After the QM training.* The participants all indicated that they had a good experience with their particular kind of training:

    "My experience with the self-training process was good. The rubric and expectations were direct and concise. I felt I could adequately assess my course through the established criteria." (self-training)

    "I enjoyed the short instruction process ... I appreciated the consideration of our time limitations and the amount of time for the initial training was about right." (short training)

    "I found the training very informative. Since I had only taught one online course prior to participating in the training, it was very helpful for me. [There were] lots of things that I had not thought of related to online course set up." (long training)

    The facilitator and the handouts were considered strengths of the training:

    "The materials that they gave [us] *were excellent resources to refer to during the review process."* (short training)

    "I believe I was prepared to do the self-review of my course. I referred to the training manual standards and annotations a lot during the review – found the information in the annotations very helpful as I worked on each standard." (long training)

    Over half of the participants mentioned that practicing the assessment of QM standards on their own courses would be a valuable addition to the training sessions:

    "I would have liked to have had the opportunity to walk through some of the questions specific to my courses with one of the instructors." (short training)

    "I think the training would have meant more to me or been more applicable if I had to review my own course while doing the training. It would have solidified the concepts from my own point of view (my own course) first and then I could apply what I learned." (long training)

In addition, half the participants noted there was a lack of time to *"review and internalize the training information after the training"* due to *"other things occurring"*:

> *"My own personal limitation was just time to spend on the project and the mental energy in addition to the other things going on simultaneously during the training time period."* (long training)

- *After self-review.* Participants all agreed that applying the rubric to their own courses was not difficult:

> *"It was easy to see the flaws in my own course."*

> *"I found that a lot of the standards were not met in my course and this was fairly easy to determine."*

> *"It was better to review my own course as I knew where things were and were supposed to be. It was a bit harder to be more objective."*

Quantitative data supported this perspective that the rubric was easy to use. Five of the six participants agreed with certified peer reviewers the majority of the time about which standards were present and absent in their courses (range of 60.0% to 72.5% agreement).

Finally, lack of time was mentioned by everyone who commented in this forum:

> *"I think I sort of rushed through the review."*

> *"The challenge was having enough time to really critically look at the course shell."*

> *"Time is an element as well – I think I probably rushed a bit more reviewing my own course."*

- *After receiving peer feedback.* Participants indicated that the detailed, comprehensive comments were carefully done, helpful, and thought provoking:

> *"The peer reviewers' comments reflected a conscientious examination of my course. I was pleased with the level of detail provided, and felt that the course was justly examined. The comments acknowledged strengths as well as areas which required further clarification."*

> *"Peer reviewers' comments were helpful, comprehensive. Suggested things that I did not think of."*

> *"It was helpful for the peer reviewer to make specific comments."*

> *"I felt like the peer reviewer took time to look at my course and was serious about providing some sound feedback."*

While the majority of participants observed similarities between their self-review and the peer review, all mentioned noting differences in the way peer reviewers scored their courses on various standards, mostly somewhat lower. Accepting that kind of feedback was a challenge. Overall, participants found the experience of peer review to be positive:

> *"It is difficult for me to be objective with my own course because I am so close to the content."*

> *"It was immensely helpful to have another instructor examine the course for items that perhaps I thought were there, but perhaps [were] not as clear or well defined as they could be."*

> *"Some people are just more lenient on grading and giving the 'benefit of the doubt' while others are much more strict. This is where I saw a lot of variability in how courses were scored."*

> *"I was really quite offended when my peer reviewer gave my course such a poor score. But ... after further review and reading her comments, I saw that she was really right about all the QM points. It was her extensive comments that made the changes a bit*

easier, although it was actually quite a painful process changing everything to meet the standards."

- *After updating their courses.* Participants employed a few other resources to help make changes in addition to the self and peer reviews and QM materials. Having been instructed not to consult the instructional designers that are available for course development, none of the participants did that, but two did consult colleagues, especially to coordinate with other sections of their course. Student feedback was mentioned by two other participants:

    "I collaborated with another faculty member who is teaching another section of the course. We wanted to have the two sections offered in a similar fashion (assignments, content, etc.) so we worked together on the updates."

    "I predominately used the self and peer review to update the course. I did consult one of my colleagues regarding my unit and course objectives before I submitted the final changes."

    "I used primarily self and peer review comments along with the QM materials that were provided ... I really did rely on the materials and the rubric provided more than anything."

Two participants decided to retain some aspects of the course that the peer reviewer had indicated should be changed; overall, participants' responses to the process of updating their courses were positive. Time to do the work was a common problem for participants:

    "The reviewer felt that there were too many unit objectives ... I disagreed and did not remove or change any of the unit objectives."

    "I had already identified all the areas that needed to be changed so the peer review just validated what I had already discovered/decided to change."

    "After I started implementing all of the suggested changes I began to realize how valuable the process of updating or implementing QM concepts will be to my students, and that is my main goal – to have the online experience a great one for students that has as little confusion as possible."

    "I found it difficult to find time to do this until the end of the semester as I prepare for the next semester. Time is the biggest issue."

    "Time consuming but necessary! Through the process, I found ways that I could enhance the curriculum of the course and took time to review learning activities and assignments and tweak them to make the online learning experience more enjoyable."

    "The process was fairly time consuming, but it has been a positive learning experience, and I now have a better sense for how I can improve the other courses I teach online."

- *After receiving feedback from the review team.* Participants noted that some of the suggestions made by the QM reviewers were similar to those of the peer reviewers and some related to aspects of the course that had not been mentioned before:

    "[The comments] seemed to be different in many ways, which is why the course still did not pass the final review. I revised the course per my peer reviewer's comments and still did not pass the QM master review."

    "It seemed that the external reviewers did not really understand clinical courses (the content expert did). So some of the comments from the external reviewers seem to be due to a lack of understanding of clinical rather than the set up of the course."

    "Yes, there were suggestions that were new. In particular, related to using color. I use a lot of color and it was suggested that I use black print throughout (or a different color) consistently and use color 'to mean something' ... [I have heard the opposite at a conference.] *I will examine this more closely. I realize it is perhaps a minor issue and it interests me (theoretically)."*

Finally, several participants said that the official QM review had been a positive experience, while two participants felt that the feedback from QM reviewers was less supportive and more critical in tone than the peer reviewers' comments:

> "I absolutely loved it! I would have liked more time to discuss this and learn. The feedback was valuable. More importantly, the process encouraged me to review and critique my own online courses."

> "In the final review of my course, the external reviewer was quite critical, using the kind of terminology that I was told not to use."

> "I think that working with seasoned educators, the QM staff/reviewers need to be very delicate in their suggestions ... maybe find some way to comment on how far the course has come and give the faculty some credit for the time and extensive efforts they have put into improving their courses instead of just providing criticism."

- *At the end of the study.* Faculty volunteers were asked about the experience of participating in the study. There were no comments from the self-training group. Those in the short training group indicated that they thought that their kind of QM training was adequate preparation for self-review while those in the long training group mentioned that additional practice would be helpful:

> "Yes [the kind of training] *was appropriate. It gave me an opportunity to review the standards with others. Then apply it and finally revise according to the guidelines."* (short training)

> "I believe the training was good, I think I just need more practice in looking at a variety of courses and how they meet or do not meet the standards." (long training)

> "Even though I had the long training, I think that I could have had more education (even though I didn't have the time). The reviewers found things that I didn't, which speaks to my lack of ability to review effectively ... however, it may come with more exposure to the guidelines as well."

In relation to making the needed changes in their courses, those in the short training group felt confident in their ability, while those in the long training group were less sure of their ability to make changes after having received what they perceived to be critical reviews from the QM teams:

> "I feel I have the ability to take my other online courses and update them to meet QM standards. I do not feel that I need any additional training."

> "Not sure – I think I can make the changes but not sure they will be adequate for the QM review unless the reviewers have a better understanding of clinical courses."

> "I think I would be able to make changes in another course, but in my own I wasn't very proficient per my peer reviewers comments after I thought I had fixed things."

A seventh instructor who dropped out part way cited the need for help making changes, which was not built into the study:

> "The short training process was very useful. However, it would have been much more beneficial to then work on my learning objectives with expert assistance. I was able to do the self-review without problem. Correcting the issues I discovered presented a problem in the assessment development."

Those in the self-training group sought help from instructional designers after participation in the study was complete. It is not known whether these instructors requested help with essential standards or less important ones; there were scattered scores of zero for all three kinds of standards on their rubric results.

The quantitative data showed that five of the six participants had improved aspects of their courses as reflected in higher scores on the QM team rubric than had been received on the peer-review rubric. The increases ranged from 10.6% (on the course that scored highest on peer review and had little to improve) to 60.0% (on the course with the lowest score on peer review, the most to improve, and the most actual improvement, achieving certification).

Overall, participants were glad to have been in the study but found it difficult to complete the needed activities:

> *"I have found this experience extremely valuable. The biggest problem (challenge) was time to work on the course ... It worked best for me to do the revisions when I could apply the revisions to something I needed to do. As I prepared for the 'next' course, I was able to spend the necessary time mak*[ing] *revisions."*

> *"If I had more time, I probably could have done a better job in updating the course ... However, I did learn a lot and believe that the course is much better put together than it was prior to this experience."*

> *"I found that I was always behind in terms of the deadlines and the entire process took a lot more time than I anticipated. If I had more time, I probably could have done a better job in updating the course."*

Participants had a few final suggestions to make related to the QM peer-review process:

> *"The biggest help was having someone to talk with after the review. More of this would be great!"*

> *"A blog or focus group would be a great way to offer support and guidance to each other re: online education. This would be helpful to me since most of what I do (and like) is online teaching."*

> *"I personally would like to see a consideration for enrollment when evaluating the delivery of course materials within the rubric ...* [I teach two sections with 60+ students every semester]. *The numbers affect 'how' I manage certain aspects of the course, but this did not seem to be a consideration in the review."*

Participants also offered comments that captured the overall benefits of peer review for the sake of improving their online courses:

> *"This study allowed me ... to see the flaws in my course and enhance the educational experience for my students."*

> *"I liked the process and found it extremely helpful. It pushed me to try to raise my standards."*

> *"The QM process supports good pedagogy, and that is truly the aim!"*

> *"The process was fairly time-consuming but it has been a positive learning experience, and I now have a better sense for how I can improve the other courses I teach online."*

Looking at Self, Peer, and QM Reviews

- *Differences associated with length of training.* In order to assess for differences among results of course reviews performed by participants who experienced varied kinds of training, participant and peer review scores and recommendations for each course were examined. The degree of similarity between scores and recommendations from a participant and a peer reviewer as they evaluated the same course was interpreted to reflect knowledge gained by the participant through their training in the use of the QM rubric. The percentage of agreement between instructor and peer reviewer on the presence or absence of standards was calculated and averaged for each pair of participants, thus representing the three kinds of training. Differences might be related to training, previous experience teaching online, or other factors; the sample was too small for statistical analysis so simple description is provided. There was a difference between the scores of faculty participants in the self-training and long training groups that trended in the direction of increased agreement being associated with longer training (Table 2). Scores from self-training participants were the same as those of their peer reviewers on 60.0% and 62.5% of the 40 standards for a mean of 61.3%. Scores from long training participants agreed with those of their peer reviewers on 67.5% and 72.5% of the 40 standards for a mean of 70.0%. The scores from short-training participants were quite disparate, containing both the lowest and one of the highest rates of agreement on QM rubric scores. The mean did not fall between the means of the self and long training groups (72.5% and 33.3% agreement for a mean of 52.9%).

Except for the one low percentage of agreement (33.3%), the others all ranged from 60.0% to 72.5%. If the assumptions made regarding interpretation of the data are correct, some kind of training about the QM rubric may have enabled most faculty members to judge accurately roughly two thirds of the 40 standards.

Table 2. *Comparison of self and peer scoring of QM rubric*

Participant	Percentage match of self and peer review scores	Mean percentage match per training group	Comments
1 – self-training	60.0%	61.3%	
2 – self-training	62.5%		
3 – short training	72.3%	52.9%	Most experienced online instructor
4 – short training	33.5%		Least experienced online instructor
5 – long training	67.5%	70.0%	
6 – long training	72.5%		

Note. Due to the small sample, information matching participants and their courses cannot be shared in detail because it could reveal participant identity.

It is interesting that comments from the participants did not reflect a universal feeling of confidence that they were able to accurately use the rubric, though numerical evidence of what might be an effect of more time in training suggests that the long training group was more accurate than the other two. Those in the long training group indicated that they wanted more practice (e.g., *"I believe the training was good, I think I just need more practice in looking at a variety of courses and how they meet or do not meet the standards."*). Short training participants did not hesitate about their skill (e.g., *"I feel I would have met standards had I not had the peer review or the final course review as no major changes were made based on the feedback given in these."*). Self-review participants did not respond in this discussion forum about adequacy of training for reviewing their courses.

Another characteristic of the data was the difference between actual scores assigned by self and peer reviewers. In five of six cases, the self-review scores were higher than the peer-review scores. The differences ranged from 13 to 38 fewer points (15.3% to 44.7%), and there was no apparent pattern related to length of training.

The data were also examined for patterns related to specific standards receiving scores of zero, indicating that changes need to be made in the course. None of the 40 standards stood out as more frequently receiving zeros than others, and only category #6 on *Course Technology* had zeros more than a third of the time (four of six courses). Of the seven standards in #6, two had zeros in half the courses (6.2 – Tools and media support engagement and guide student to be an active learner, and 6.7 – Course design takes full advantage of available media) and one had zeros in two thirds of the courses (6.4 – Students have ready access to technologies required in the course). One of the participants commented several times in discussions that course technology was an area about which more information needed to be shared in the training session. No pattern related to length of training, experience teaching online, or other characteristics was observed in the scoring of QM rubrics.

Agreement on recommendations written by faculty participants and peer reviewers was also examined. When a score of zero is assigned, QM training encourages reviewers to write supportive comments about how an instructor has partially met the standard and to provide suggestions for implementing improvements. There was no pattern observed related to length of training. For four of the six courses, peer reviewers wrote more recommendations than the participants did as part of self-review.

- *Faculty ability to incorporate suggested changes.* In order to assess the ability of faculty participants to take the results of self and peer reviews and improve their online courses without additional training or help with instructional design, the scores and recommendations from peer reviews and official QM reviews were examined for differences. An increase in scores from peer review to QM review was interpreted to reflect successful course improvement (Table 3). There

was an increase in scores for five of the six courses, ranging from 10.6% to 67.5% of the 40 standards, and five of the six achieved the necessary score of 72 points or more for course certification. One of those five met the other criterion of exhibiting the presence of all of the three-point Essential standards and qualified for certification during the study, while the other four courses needed only one to three standards to be updated further (which was accomplished after the study). There was no observable pattern associated with length of training. The only course that showed a decrease in scores from peer to QM review was the hybrid clinical course. Scores and recommendations from two of the QM reviewers may have reflected an unfortunate characteristic of the rubric: there is no flexibility in judging the need to include all the standards, so all courses are expected to have components that fit a traditional three-credit content-based course. A course that is different thus earns scores of zero even though absence of certain standards would be the appropriate condition from the instructor's viewpoint. For example, there is no content presented in this online part of the clinical course, so weekly readings and related objectives would not be expected.

Table 3. *Improvement in rubric scores after course revision*

Participant	Percentage increase in score from peer to QM review	Comments
1 – self-training	22.3%	
2 – self-training	17.7%	
3 – short training	10.6%	Highest score at first, less to change
4 – short training	67.0%	Big change, achieved QM certification
5 – long training	-15.3% (decrease)	Atypical, hybrid clinical course; did not fit rubric well
6 – long training	25.9%	

Another characteristic of the data that was examined was zero scores given by peer and QM reviewers. Three combinations were of interest: zero to full points, zero to zero points, and full to zero points. The case of a standard receiving zero points upon peer review then full points upon QM review was interpreted to mean that the course had been improved sufficiently to meet the standard. This was the desired outcome, and in five of six courses this was the most common situation that involved scores of zero (32% to 89% of the standards moved from zero to full points in various courses). The second situation involved zero scores assigned by both peer and QM reviewers and was interpreted to mean that the participant either did not make sufficient changes, decided not to make the suggested changes, or experienced the results of a disagreement in the assessments of peer and QM reviewers. Most standards did not have repeated zeros; four courses had only one to three of the eight standards with repeated zeros. The third situation was a change from full points given by the peer reviewer to zero points from the QM team. This situation was interpreted to mean that the participant either made changes that impaired meeting the standard or experienced the results of disagreement in the assessments of peer and QM reviewers. Again, four courses had only one to three standards in the full to zero points situation. There was no observable pattern that corresponded to length of training.

Discussion

Using the Rubric and Updating Courses

Agreement between participants and their peer reviewers on scores from the QM rubric was interpreted to mean that the faculty participants had learned to use the rubric accurately. There was a difference between groups – a positive trend toward more agreement as participants had more training time – but statistical testing was not possible. Comments from participants indicated that confidence in their ability to accurately use the rubric for course evaluation did not match the trend in scores. Perhaps the long training, which involved three more hours of practice time, made the participants more aware of how judgments are made when using the rubric, or perhaps the critical tone of feedback from some QM reviewers caused participants to doubt their ability to use the rubric. Nevertheless, with some kind of training, participants were able to agree with certified peer reviewers on the status of roughly two thirds of the standards in their courses.

If the suggested underlying pattern of a positive relationship between length of training and ability to accurately self-review courses proves reliable upon further research, it appears that self-training (61.3% agreement) is nearly as good as long training (70.0% agreement) in terms of helping faculty members learn to assess their courses for the meeting of QM standards. The issue will be what degree of accuracy in online course review fits with a given school's goals for participating in the QM program. If each course is going to be improved until it can achieve course certification, some kind of universal training for online faculty members might be a basic first step in that process, followed by peer course review and assistance to meet sufficient standards. If a broad, voluntary approach to faculty development related to improvement of online courses is chosen, without a push for course certification, some kind of faculty training might be made available to facilitate online course improvement in a general way, with expert assistance available for faculty who might seek it. Assisting instructors to learn enough to accurately identify the presence of roughly two thirds of the 40 standards is a good place to begin in either case.

Some instructors may already know enough about educational principles and good online practices to take advantage of minimalist training approaches such as the self-training used in this study. Other instructors have more need for classes and additional expert assistance. Perhaps a screening test could be developed so faculty members can be guided to training options, which will help use everyone's time and resources efficiently.

The difference between self and peer scores, for which most of the self-review scores were higher than peer review scores, is perhaps related to the spontaneous comments of the majority of the participants about the *"natural leniency"* of some reviewers. One participant attributed this to personality – having or not having the tendency to *"give people the benefit of the doubt."* The QM experts associated with the study team thought this was more likely to be related to beginning development of an ability to estimate the 85% level of presence that is to be identified when assessing a given standard. The importance of the higher self-review scores is that instructors could think they have met more standards than a certified reviewer would identify, and would not necessarily continue to improve those standards they thought they had met, when in reality there is more improvement needed than they realize. If the figures of 61.3% for self-training and 70.0% for long training are found to be reliable, there is a gap of 30% to 40% in the rate of identification of needed changes by instructors versus peer reviewers. Depending on the approach of the school to the use of QM, that could pose a significant and unacceptable difference.

Related to differences in the written recommendations generated by course reviews, the peer reviewers demonstrated that they were more likely to provide recommendations than the participants themselves. Since all of the local peer reviewers had completed several levels of QM training in order to become certified reviewers, it makes sense that they understood and took seriously their responsibility to write comments. The participants' training options had not covered this aspect to the same degree, and it may have felt awkward to write recommendations to one's self, as their course self-evaluations would have them do.

As to participants' ability to improve online courses, an increase in scores from peer review to QM review was present and was interpreted to reflect successful efforts to update their courses. There was an increase in total scores for five of the six courses, ranging from 10.6% of the 40 standards (this course had the highest score upon initial peer review so there was less to change) to 67.5% (this course had the lowest score so it had the most to change, which was done, and certification was achieved). It was apparent that most instructors could accomplish a good deal on their own but at one point or another, in discussions or informally, all of the participants inquired about getting some help with some parts of their courses that needed improvement. They were asked to wait until after they completed the study, and several participants did access help at that point. This finding supports the need for resource people to be available for faculty as they work on the instructional design of their courses. The help of instructional designers is better utilized now than in the past, but it would not be surprising if many faculty members still see the design and updating of their online courses to be their responsibility exclusively. Effecting change in faculty culture takes time and effort, though the outcome of improved course design that supports student learning online would be a worthwhile result. Funding is a variable that intersects with faculty ability to access help with instructional design. The issue of finding or creating financial resources to provide instructional designers and other sorts of personnel needed to support online programs is an increasingly serious and difficult one in higher education. Determining cost effective and instructionally efficient faculty training methods will be important as part of managing the financial aspects of faculty development.

There was notable variation between peer reviewers in the number and length of the supportive comments and suggestions for improvement that they provided as part of the recommendations after course review. Both extremes – few or no comments vs. extensive, detailed comments – have drawbacks. If assigning a zero for a given standard, reviewers are obligated to provide information to guide the instructor. Too many suggestions can be perceived as controlling, and referring to other courses was not welcomed. There was also variation in the assessment of particular standards, leading to zero scores upon which the peer and QM reviewers disagreed. The QM expert on our study team noted that the potential for disagreement between reviewers is one reason for having three reviewers on an official team, to "break a tie." Being on the receiving end of inconsistent evaluations in the official QM reviews was difficult for participants since they had based course improvements on peer reviews which then weren't sufficient for the QM reviews.

Comments from Participants

Several themes were clearly evident. First, the whole process of QM course review and revision was valuable to faculty members. It supported their personal goals for improving the educational experiences of their online students. Participants found the QM materials helpful, the training sessions useful, and the peer and QM comments detailed and instructive. They had several suggestions for improvement of the process such as incorporating faculty members' own courses into training sessions, developing flexibility in expectations related to the demands of course enrollment, and maintaining a positive tone when providing feedback as a peer reviewer. These strong statements in favor of a process to assist faculty members in improving their online courses for the sake of their students is important; this finding provides support for continued efforts by faculty and school administration to meet this goal.

Second, the overriding challenge for participants related to both participating in the study and attempting to improve their online courses using the QM program was finding the time to work on them given faculty workload, personal responsibilities, and life events. While by no means new, the issue of limited faculty time clearly makes it difficult to accomplish the sort of in-depth course review and revision that is periodically necessary. Suggestions from participants for dealing with that issue included timing course improvement efforts to occur when faculty would be updating courses for the next term and providing online blogs or discussions to support faculty asynchronously. Participants did not suggest that course review should be handed over to anyone else to do; faculty ownership for quality and updating of courses was strong among these faculty members.

Third, several participants mentioned the need for face-to-face help from colleagues and instructional designers, and indeed sought that out. Although there has been some change in recent years in terms of helping new faculty members learn best practices as they begin their teaching careers, it remains likely that most faculty members in higher education do not have much preparation for teaching in their disciplines. On-the-job training can work well, but faculty members must somehow make time to pay attention to information about selecting content, writing objectives, aligning them with readings, managing discussions, designing other learning exercises, and devising assessments, as well as understanding the technical aspects and other principles related more specifically to online learning. Even experienced educators may struggle with all of this when designing or updating a course, so there is a valuable and necessary place in each academic institution for specialists in instructional design. The implication of this finding for administrators is to keep searching for ways to provide faculty development and staff support related to online course improvement. Clearly there was a need to have experts available to help when instructors got stuck in their efforts to understand and work to improve a given standard.

Limitations and Recommendations

There were several limitations related to this study. The small number of faculty participants, differences in their rate of progress through the stages of the study, and lack of time for the study activities hampered full participation in the asynchronous discussions and rushed self-evaluation of online courses. Despite these challenges, saturation of qualitative data was noted.

The major limitation of this study relates to how to use the findings given its small sample and use of simple descriptive statistics. The results should not be generalized but can be considered in light of local conditions in other educational institutional settings and used as a starting point for discussions about the need for instructional design support for faculty and the goals for implementing QM or another approach for improvement of online course quality.

While it would be possible for the study to be replicated, perhaps a better recommendation for future research would be the design of a mixed-methods multi-site project based on the descriptions generated by this study. One could find a set of educational institutions that are interested in trying the self-training approach to faculty development and improvement of online courses. A longitudinal research project could follow that process in multiple locations, studying and comparing the experience of faculty members, testing knowledge acquisition about use of the rubric, and examining the kinds of course changes and associated student performance that result. In contrast, it might be possible to find schools interested in supporting faculty development and improvement of online courses leading toward universal QM course certification. Tracking expenditures and testing for improved satisfaction and enhanced academic performance of students in the included courses would be valuable parts of such studies.

There are several other variables that are part of this complex situation and could be studied. Examples are faculty culture, faculty motivation to update courses, and current skills of faculty members for instructional design work such as that done in this study. Another approach is strictly financial: analyze costs related to the time spent by faculty members on course review and improvement compared with predicted costs for turning that activity mostly or entirely over to instructional designers.

If subsequent research supports the observation from this study that a self-study or shortened format for QM training provides sufficient faculty learning related to use of the rubric, a source of savings in time and costs for both faculty and educational support services staff has been identified.

Conclusion

The main findings of this study are as follows:

1) Short-duration training (none to six hours of class) allowed online instructors to learn to use the QM rubric accurately enough to match about two thirds of the scoring of certified peer reviewers;

2) Instructors used the results of self and peer reviews to update many aspects of their online courses without additional help from instructional designers or other experts;

3) Instructors reported they needed to have access to instructional designers or other experts for help with some aspects of their courses that did not meet the standards;

4) Instructors reported they were eager to learn how to improve the design of their online courses for the sake of student learning experiences;

5) Instructors reported that heavy workloads made it difficult to find time for course evaluation and revision;

6) Challenges inherent in the process for QM scoring and writing recommendations may influence instructors' perception of receiving fair and respectful feedback.

The main recommendation from this study is that faculty open a discussion in their departments or schools about how to improve the instructional design of their online courses. The following questions should be asked:

- What kind of system or standards should be used?

- Who should be responsible for organizing and carrying out the reviews?

- Who should be responsible for implementing needed course changes?

- How often should course review be completed?

- What kind of faculty and staff training would be necessary and sufficient?

- What other kinds of support would best achieve the goal of improved student learning (instructional designers in particular)?

- What combination of the factors above would be financially feasible?

References

Adams, C. L., Rust, D. Z., & Brinthaupt, T. M. (2011). Evolution of a peer review and evaluation program for online course development. In J. E. Miller & J. E. Groccia (Eds.), *To improve the academy: Resources for faculty, instructional, and organizational development* (Vol. 29, pp. 173-186). San Francisco, CA: Jossey-Bass.

Cobb, K. L., Billings, D. M., Mays, R. M., & Canty-Mitchell, J. (2001). Peer review of teaching in web-based courses in nursing. *Nurse Educator, 26*(6), 274-279. doi:10.1097/00006223-200111000-00012

Cohen, P. A., & McKeachie, W. J. (1980). The role of colleagues in the evaluation of college teaching. *Improving College and University Teaching, 28*(4), 147-154. doi:10.1080/00193089.1980.10533656

Creswell, J. W., & Plano Clark, V. L. (2010). *Designing and conducting mixed methods research* (2nd ed.). Newbury Park, CA: Sage.

Gaytan, J., & McEwen, B. C. (2007). Effective online instructional and assessment strategies. *The American Journal of Distance Education, 21*(3), 117-132. doi:10.1080/08923640701341653

Gunawardena, C. N., Ortegano-Layne, L., Carabajal, K., Frechette, C., Lindemann, K., & Jennings, B. (2006). New model, new strategies: Instructional design for building online wisdom communities. *Distance Education, 27*(2), 217-232. doi:10.1080/01587910600789613

Little, B. B. (2009). The use of standards for peer review of online nursing courses: A pilot study. *Journal of Nursing Education, 48*(7), 411-415. doi:10.3928/01484834-20090615-10

Legon, R. (2006). *Comparison of the Quality Matters rubric to accreditation standards for distance learning*. Baltimore, MD: Quality Matters. Retrieved from https://confluence.delhi.edu/download/attachments/74055682/Comparison+of+the+Quality+Matters+Rubric+-+Summary.pdf

MarylandOnline. (2008). *The Quality Matters program rubric*. Retrieved September 13, 2010, from http://www.qmprogram.org/rubric/ (archived at http://web.archive.org/web/20100911002421/http://www.qmprogram.org/rubric/)

MarylandOnline. (2013a). *About us*. Retrieved from http://www.qualitymatters.org/about/

MarylandOnline. (2013b). *Higher ed program > Rubric*. Retrieved from http://www.qualitymatters.org/rubric/

McNaught, C. (2001). Quality assurance for online courses: From policy to process to improvement? In G. E. Kennedy, M. J. Keppell, C. McNaught, & T. Petrovic (Eds.), *Meeting at the crossroads: Proceedings of the 18th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 435-442). Melbourne, Australia: Biomedical Multimedia Unit, The University of Melbourne. Retrieved from http://www.ascilite.org.au/conferences/melbourne01/pdf/papers/mcnaughtc.pdf

Millis, B. J. (2006). Peer observations as a catalyst for faculty development. In P. Seldin (Ed.), *Evaluating faculty performance: A practical guide to assessing teaching, research, and service* (pp. 82-95). Bolton, MA: Anker.

Morehead, J. W., & Shedd, P. J. (1997). Utilizing summative evaluation through external peer review of teaching. *Innovative Higher Education, 22*(1), 37-44. doi:10.1023/A:1025199425293

Nelson, M., & Van Leeuwen, P. (2008). Peer review of online courses. In C. J. Bonk, M. M. Lee, & T. Reynolds (Eds.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2008* (pp. 442-446). Chesapeake, VA: Association for the Advancement of Computing in Education. Available from EdITLib Digital Library. (29643)

Puzziferro, M., & Shelton, K. (2009). Supporting online faculty – revisiting the seven principles (a few years later). *Online Journal of Distance Learning Administration, 3*(3). Retrieved from http://www2.westga.edu/~distance/ojdla/fall123/puzziferro123.html

Quality Matters. (2013). *Subscriber institutions by country*. Retrieved from http://www.qmprogram.org/qmresources/subscriptions/subscribers.cfm

Ross, K. R., Batzer, L., & Bennington, E. (2002). Quality assurance for distance education: A faculty peer review process. *TechTrends, 46*(5), 48-52. doi:10.1007/BF02818309

Russell, T. L. (2001). *The no significant difference phenomenon: A comparative research annotated bibliography on technology for distance education* (5th ed.). Montgomery, AL: International Distance Education Certification Center.

Western Interstate Commission for Higher Education Cooperative for Educational Technologies. (2010). *No significant difference*. Retrieved from http://www.nosignificantdifference.org/

Wiesenberg, F., & Stacey, E. (2005). Reflections on teaching and learning online: Quality program design, delivery and support issues from a cross-global perspective. *Distance Education, 26*(3), 385-404. doi:10.1080/01587910500291496

Acknowledgement