March 2014

Continuous Improvement of the QM Rubric and Review Processes: Scholarship of Integration and Application

Kay Shattuck
The Quality Matters Program

Whitney Alicia Zimmerman
The Pennsylvania State University

Deborah Adair
The Quality Matters Program

Follow this and additional works at: http://digitalcommons.apus.edu/internetlearning

Part of the Educational Assessment, Evaluation, and Research Commons, and the Online and Distance Education Commons

Continuous Improvement of the QM Rubric and Review Processes: Scholarship of Integration and Application

Kay Shattuck[A], Whitney Alicia Zimmerman[B], Deborah Adair[C]

Quality Matters (QM) is a faculty-centered, peer review process that is designed to certify the quality of online and blended courses. QM is a leader in quality assurance for online education and has received national recognition for its scalable, peer-based approach and continuous improvement in online education and student learning. Regular, robust review and refreshment of the QM RubricTM and processes keep them current, practical, and applicable across academic disciplines and academic levels. The review ensures validity in the set of quality standards that make up the Rubric. An overview of the regular review of the QM Rubric and process, as well as examples of the use of data to continuously improve the Rubric and process are presented. The guiding principles of QM – a process that is continuously improved upon and that is collegial and collaborative – are discussed in relationship to Boyer's scholarship of application and scholarship of integration. Glassick (2000) noted that Boyer's scholarship of overlapping discovery, integration, application, and teaching is "a hard but worthwhile task" (p. 880). This article outlines how the dynamic and rigorous processes adopted by QM continue to take on that worthwhile task.

Keywords: Quality Matters, course design, professional development, continuous improvement, quality assurance, rater agreement

Introduction and Background

The Quality Matters (QM) Program was initially developed under a 2003–2006 Department of Education Fund for the Improvement of Post-Secondary Education (FIPSE) grant. The grant, awarded to the not-for-profit consortium, MarylandOnline, was for the development of a replicable quality assurance program focused on faculty peer review and improvements to the design of online courses. During the grant period, a community of practice within Maryland researched, developed, imple- mented, and disseminated a set of quality benchmarks (standards) (Shattuck, 2007), as well as a rigorous peer review process to improve student learning in online courses. In their wisdom, the developers of the QM program recognized that providing an in- strument (a Rubric) and a process for using this Rubric would not be enough. Drawing from their own experiences as members of a community of practice that worked to- gether for many years to solve the common issue of improving online course designs (Cervero & Wilson, 1994; Lave & Wenger, 1991; Schön, 1983; Cousin & Deepwell,

---

[A] Director of Research, Quality Matters Program

[B] Doctoral Student, The Pennsylvania State University

[C] Managing Director and Chief Planning Officer, Quality Matters Program

2005; Guldberg & Pilkington, 2006), they included required credentialing of specific competencies in the use of the Rubric and in an understanding of the application of the QM guiding principles of being collaborative, collegial, continuous, and centered in academic foundations around student learning.

Quality Matters is a program that subscribing educational institutions use within the cadre of other components necessary to assure quality in their online learning programs. While the QM Rubric[2] is focused on the design of online and blended courses, the QM process was developed with the awareness that it impacts faculty readiness through the QM professional training program (emphasizing pedagogical underpinning of course design), as well as the benefits of collegial interactions across academic disciplines and educational institutions. Other factors affecting course quality include course delivery (teaching), course content, course delivery system, institutional infrastructure, faculty training/readiness, and student readiness/engagement. The importance of other components in an institution's quality assurance commitment to online education is acknowledged within the QM standards.

Quality Matters is a faculty-centered, peer review process that is designed to certify the quality of online and blended courses. QM is a leader in quality assurance for online education and has received national recognition for its scalable, peer-based approach and continuous improvement in online education and student learning. As of the winter of 2013, there are 825 subscribing educational institutions and 160 individual subscribers; 3,998 courses have been formally peer reviewed; and 28,756 online educators have successfully completed QM professional development courses.

In this article, the QM guiding principles – a process that is continuously improved upon and that is collegial and collaborative – are discussed in relationship to Boyer's scholarship of application and scholarship of integration. An overview of the regular review of the QM Rubric and process, as well as examples of the use of data to continuously improve the Rubric and process are presented.

Scholarships of Application and Integration

While the construct of CoP (community of practice) (Shattuck, 2007) is useful in understanding the developmental phases of the QM program, the past decade can be described as an evolving practice of Boyer's (1990) scholarships of application and integration. In a seminal publication of The Carnegie Foundation for the Advancement of Teaching, Scholarship Reconsidered, Ernest Boyer challenged higher education to move beyond "teaching versus research" (p. 16) and for faculty to take on a scholarly approach to teaching by rigorous study of teaching in ways that are collaborative and connect theory with the realities of teaching. The term "the scholarship of teaching and learning[1] (SoLT)" is becoming an increasingly familiar concept in higher education (Hutchings, Huber, & Ciccone, 2011). Lesser known is that Boyer suggested "four separate, yet overlapping, functions" (p. 16) of scholarship. Those are the scholarships of discovery, integration, application, and teaching, and have been applied as useful tools in defining scholarship (AACN, 1999).

- The scholarship of discovery relates to the most traditional functions of research, that is, exploration to generate new knowledge.

• The scholarship of integration is "*inter-disciplinary, interpretive, integrative*" (italics in original) (Boyer, p. 20) and about "making connections across disciplines" (p. 18).

• The scholarship of applicaton is about use of knowledge from research to improve societal problems.

• The scholarship of teaching encompasses the relationship between teacher and student in which the teacher is also a learner to improve student intellectual growth

Boyer's call "to liberate academic careers from the hegemony of published research as the dominant product and measure of scholarship" (Bernstein & Bass, 2005, para. 41) served as a "tipping point" in the century-long debate of research versus teaching (Rice, 2002, p. 7). The growing sophistication of digital technologies of the past decade introduces new formats for the production, publication, and dissemination of faculty scholarship (Bernstein & Bass, 2005; Hatch, Bass, Iiyoshi, & Mace, 2004). The scholarship of application and integration is evident in QM's research on continuous improvement. Examples described in this article are

• Regular review and refinement of the QM Rubric and peer review processes;

• Consistently rigorous applications of the QM process, which are inter-disciplinary and integrative, and provide tools and strategies for interpreting research into useable processes; and

• Statistical analyses of data gathered during the QM peer reviews which inform continuous improvement of the QM Rubric and application of research and shared online teaching/designing expertise across academic disciplines and educational institutions.

Ultimately, the scholarship of teaching is behind the QM commitment to development and dissemination of standards of quality in online course design, which is a key phase in developing strong teaching presence. The scholarship of discovery – "disciplined work that seeks to interpret, draw together, and bring new insight to bear on original research" (Boyer, p. 19) – is the focus of QM's interest in original research. This interest will be the focus for 2014-2015.

Regular Review and Refinement of the QM Rubric and Processes

The 2007 article by Shattuck describes the development of the eight general standards of quality online course design as they were (and continue to be) informed by the independent research literature and established best practices. The QM Rubric and processes are dynamically interpretive of evolving research and best practices. The plan to conduct a complete review of the QM Higher Education Rubric and peer review process was established during the grant period, and reviews have become more thorough over the past decade. The ongoing history of review and refinement of the QM Higher Education Rubric and Processes chart outlines the review process and outcomes for the past five Rubrics, from the first to the current review.

The chart outlines the continuously improving processes used by QM to ensure wide input and transparency in the refinement of the Rubric and the peer review process. Figure 1 represents the current, rigorous, and comprehensive process followed to launch each new edition of the QM Rubric. The process is undergirded by the commitment to interpret research, best practices, and teaching/designing expertise

into an applicable process that can be used across all academic disciplines. The collaboration of peer reviewers across disciplines points to Boyer's scholarships of application (practice) and integration.

Consistently Rigorous Application of the QM Peer Review Process

Following the principles of faculty-centered and continuous improvement, the QM higher education Rubric has been thoroughly reviewed and refined to ensure it remains a current and effective set of quality guidelines in online course design. It is important to recognize that while there is an openly accessible listing of QM standards, the full QM Rubric contains detailed annotations for each standard that assist in interpreting and applying standards during a course review. A course review without access to the complete QM Rubric and done by non-QM-certified reviews does not meet the rigors of the QM process. QM course reviews are conducted by a team of three certified QM Peer Reviewers (PRs) – all are active online instructors, all are currently certified as QM PRs, at least one PR is from outside the institution of the course under review, and at least one PR is a subject matter expert (SME) in the academic discipline of the course under review. Each team is led by a QM Master Reviewer (MR) who has extensive online teaching experience and in the QM review process, as well as having additional training in facilitating an inter-institutional virtual collaboration of academic peers.

Each QM PR brings at least two years of current experience teaching online. Additionally, each is required to complete rigorous QM training to become QM certified; each is subsequently added to the QM database of available PRs available to conduct QM course reviews. Each certified PR's academic discipline is included in the database. Course review teams are developed using the database of certified PRs. This ensures that at least one SME related to the course under review is included on each team. While a QM review does not evaluate the content of a course, an SME serves as a resource for others on a review team on any course design implications for a particular academic discipline. Each review team is chaired by a QM MR, an experienced reviewer with advanced training on the rubric and review process, who guides the team as needed in interpretation of the standards.

Consistent Application of QM Peer Reviews

Quality Matters is sometimes mistakenly described as a "Rubric," while in fact, it is a process of engaging online faculty who have further training in their use of a validated set of standards (encapsulated in the QM Rubric). This set of standards guides reviewers in their collaborative assessment of the design quality of a particular online course. The rigorous QM peer review process that results in courses meeting QM standards of quality (either initially or upon amendments) includes formal and informal reviews of online courses and online components of blended courses. Informal use of the QM Rubric is under the discretion of the subscribing institution. Formal course reviews are either managed by the QM program staff (QM-managed) or by certified QM representatives within a subscribing institution (subscriber-managed).

The analysis of 2008–2010 data found no difference between QM- and subscriber-managed formal course reviews in terms of total points ($t(272) = 0.831$, p $=.406$) or review statuses ($\chi^2(2) = 0.500$, p

Research & Analysis

Broad Literature Review

QM Research Review

User Surveys & Feedback

Course Review Data Analysis

Rubric Committee Selection

Nomination Letters & Application

Criteria:Experienced users, broadly representative

Legacy Committee

Adhoc "Special Topic" Committees

Professional Development (PD)

Revise Entire PD Curriculum

Prepare Facilitators

Create Rubric Update Workshop

Create Other New PD as Needed

Review Process

Revise Review Management System

Revise Course Worksheets

Revise Review Policies as Needed

Communication

Update all Online Information

Ensure all QM Role Holders are Updated

Publish New Workbooks and Rubric-Related Documents

© 2014 MarylandOnline
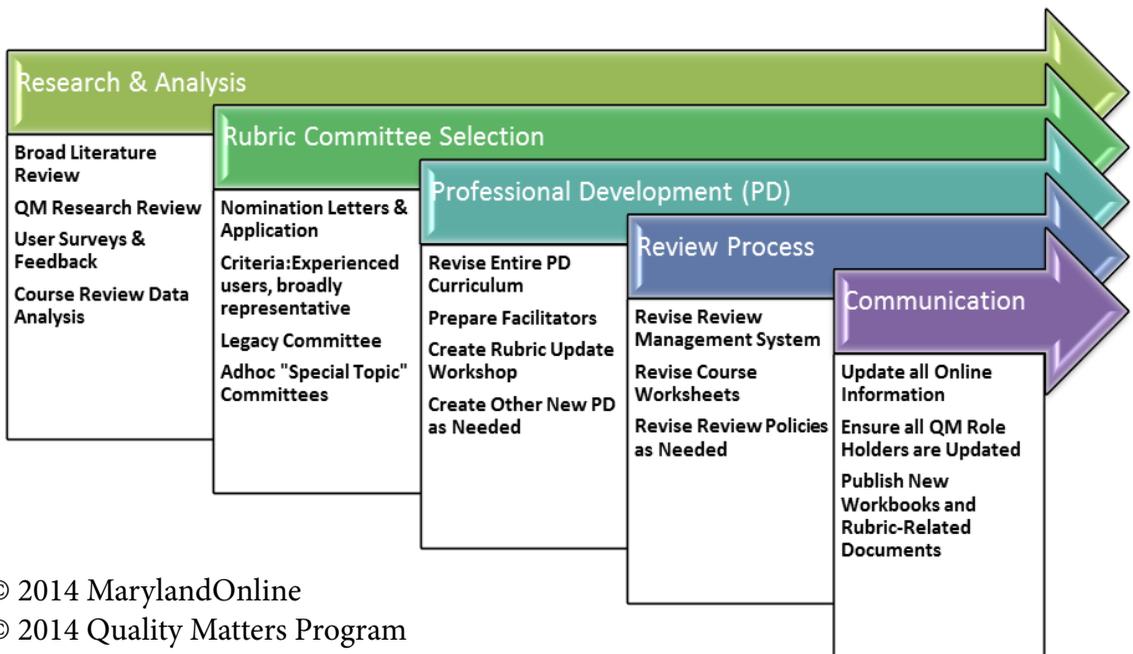© 2014 Quality Matters Program

Figure 1. Steps in QM Rubric revision process.

= .779). This attests to the level of competencies provided by QM professional development required to become a PR and to the consistency of application of the inter-institutional, faculty-focused, peer-collegial review processes.

Examination of Peer Reviews 2008–2013

Statistical analyses were conducted on data gathered from formal course reviews conducted from 2008 through 2010 (n included 434 course reviews; of those, 180 were "informally managed") and from 2011 through July 2013 ($N = 1,494$). These data were explored to identify (1) frequency of courses meeting QM standards in initial reviews and after amendments, (2) most frequently missed standards, (3) differences between courses from different academic disciplines, (4) differences between courses submitted by faculty developers/instructors with and without familiarity with the QM, and (5) proportion of inter-rater agreement by specific standards.

Rate of Courses Initially Meeting QM Standards

In the technical report for the 2008–2010 QM Rubric (Zimmerman, 2010), the results of 274 QM- and subscriber-managed course reviews were analyzed. At the time of the data analysis, 39% (105) of courses met standards in the initial course review. An additional 48% (131) did not meet standards in the initial course review but did meet standards after an amendment. The remaining 14% (38) of course reviews were considered to be in the process of amendments.

In the technical report for the 2011–2013 QM Rubric (Zimmerman, 2013), the results of 1,490 course reviews were analyzed. At the time of the data analysis, 70.5% (1,051) of courses met standards in the initial review. An additional 26.6% (397) of courses did not meet standards in the initial course review but did meet standards after an amendment. The remaining 2.8% (42) courses were pending amendment.

Explanations of the increase in courses meeting QM standards during the initial peer review include (1) more courses are being developed using the QM Rubric as a course design guide; (2) more subscribing institutions are providing informal course reviews prior to submission for a formal peer review; and (3) more faculty and design teams have acquired effective competencies in the nuances of online teaching and course design.

Most Frequently Missed Standards

For the 2008–2010 course reviews, the most frequently met standards were 6.1 and 6.5; they were both met in 96% of course reviews. The standards most frequently not met were 8.2 and 3.5; they were not met in 54% and 60% of course reviews, respectively.

For the 2011–2013 course reviews, the most frequently met standards were 6.1 and 7.2; they were both met in 95% of course reviews, respectively. The standards most frequently not met were again 8.2 and 3.5; they were not met in 58.9% and 65.3% of course reviews, respectively. Note, for the analyses of the 2011–2013 course reviews, standards met in initial reviews versus amended reviews were not distinguished between.

Frequently missed standards are reviewed carefully by the Rubric Committee, as they might indicate the need for refinement of the standard and annotation wording or need for focused QM professional development for PRs.

Differences of Review Success by Course Discipline

Analysis of courses reviewed from 2011 through July 2013 revealed that business courses tended to have the best outcomes. Business courses were most likely to meet standards in the initial review, followed by education courses. Business courses also had the highest total scores. Courses in the remaining disciplines did not significantly differ from one another.

Relationship between Faculty Developer/Instructor of Reviewed Course and Familiarity with the QM Rubric

In the analyses of the 2011–2013 course reviews, courses submitted by individuals familiar with QM had higher initial scores than courses submitted by individuals who were not familiar with QM (Mann–Whitney U ($N = 1,488$) = 43,537, p < .001). However, there were not total point differences after amendment (Mann–Whitney U ($N = 1,488$) = 61,900, $p = .108$). (The amendment phase includes interaction with the peer review team.)

The familiarity of faculty developers and instructors with the QM Rubric was examined in relation to the outcome of the initial course review and the amended course review (when needed). In the analysis of the 2011–2013 Rubric, the majority (93.3%) of individuals who submitted courses for review were familiar with the Rubric. Only 98 out of 1,492 (6.6%) of individuals stated that they were not familiar with the Rubric.

Proportion of Rater Agreement by Specific Standards

Measures of reliability are often given when discussing scores such as those assigned using the QM Rubric. The term "reliability" refers to consistency of results. Inter-rater reliability is a measure of the relationship between scores assigned by different individuals (Hogan, 2007). In its strictest sense, however, inter-rater reliability works under the assumption that reviewers are randomly selected and interchangeable (see Suen, Logan, Neisworth, & Bagnato, 1995). This assumption is not met in the QM's process in which reviewers may be selected on the basis of their previous experiences or areas of expertise. The measurement of interest concerning the QM Rubric is the proportion of reviews in which all three raters assigned the same rating to a specific standard (i.e., all three reviewers assessed a standard as met or not yet met). This is different from inter-rater reliability in that it is not an attempt at describing unsystematic variance (see Hallgren, 2012; Liao, Hunt, & Chen, 2010); its purpose is to provide an easily interpretable statistic that will allow for the comparison of specific standards for practical purposes. Thus, in the discussion of consistency of results of QM's reviews, the term proportion of rater agreement is used as it explicitly describes the analyses performed as opposed to inter-rater reliability, which it technically is not.

One of the primary purposes of analyzing proportion of rater agreement is to identify specific standards that may require attention to keep the Rubric reflective of the research and fields of practice while being workable for a team of inter-institutional, inter-disciplinarian academic peers. A specific standard for which reviewers

frequently submit different scores may lack clarity; this could result in the need for changes to the specific standard or it could signal a need for more reviewer training.

The standards with lowest rater agreement table provide an overview of the revisions made by the 2010 Rubric Committee for standards that statistically had the lowest rater agreement. The chart also provides data on the most recent data analysis and has been provided to the 2014 Rubric Committee.

Individual ratings given by a QM peer review in course reviews reflect, to at least some extent, that particular reviewer's professional/pedagogical opinion, and, therefore, may vary from the ratings of the other individual reviewers. However, markedly lower rater agreement for specific standards in the QM Rubric is a prompt to members of the Rubric Committee to focus attention on those standards during the regular review and refreshment of the QM Rubric.

Summary

Regular, robust (breadth and depth) review and refreshment of the QM Rubric and processes keep them current, practical, and applicable across academic disciplines and academic levels. The review includes interpretation of educational research, as well as an emerging emphasis on research generation. Expertise from online educators across the United States plays a critical role in the transparent, faculty-centered processes. The review ensures validity in the set of quality standards within the Rubric. Statistical analyses of data gathered from formal course reviews reveals that the peer review process has been consistently applied across review types and academic disciplines and points to the value of QM's professional development in which over 28,000 online educators have participated. The analyses also provide critical information to the Rubric Committee on the frequency of met standards and on the proportion of rater agreement by specific standards.

Glassick (2000) noted that Boyer's scholarship of overlapping discovery, integration, application, and teaching is "a hard but worthwhile task" (p. 880). This article outlines how the dynamic and rigorous processes adopted by QM continue to take on that worthwhile task. All aspects of the QM program are regularly reviewed and refreshed with and for online teaching faculty.

References

American Association of Colleges of Nursing, AACN Task Force on Defining Standards for the Scholarship of Nursing. (1999). *Defining scholarship for the discipline of nursing.* Retrieved from http://www.aacn.nche.edu/publications/position/defining-scholarship

Bernstein, D., & Bass, R. (2005). The scholarship of teaching and learning. *Academe*, 91(4), 37–43.

Cervero, R. M., & Wilson, A. L. (1994). The politics of responsibility: A theory of program planning practice for adult education. *Adult Education Quarterly*, 45(1), 249–268.

Cousin, G., & Deepwell, F. (2005). Designs for network learning: A community of practice perspective. *Studies in Higher education*, 30 (1), 57–66.

Glassick, C. E. (2000). Boyer's expanded definitions of scholarship, the standards for assessing scholarship, and the elusiveness of the scholarship of teaching. *Academic*

Medicine, 75(9), 877–880. Retrieved from http://www.academicpeds.org/events/assets/Glassick%20article.pdf

Guldberg, K., & Pilkington, R. (2006). A community of practice approach to the development of non-traditional learners through networked learning. *Journal of Computer Assisted Learning*, 22, 159–171.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23–34.

Hatch, T., Bass, R., Iiyoshi, T, & Mace, D. (2004). Building knowledge for teaching and learning. *Change*, 36(5), 42–49.

Hogan, T. P. (2007). *Psychological testing: A practical introduction*. Hoboken, NJ: John Wiley & Sons, Inc.

Hutchings, P., Huber, M. T., & Ciccone, A. (2011). *The scholarship of teaching and teaching reconsidered: Institutional integration and impact*. Hoboken, NJ: Wiley.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Liao, S. C., Hunt, E. A., & Chen, W. (2010). Comparison between inter-rater reliability and inter-rater agreement in performance assessment. *Annals of the Academy of Medicine*, 39(8), 613–618.

O'Banion, T. (1997). *A learning college for the 21st century*. Phoenix AZ: Oryx Press.

Rice, R. E. (2002). Beyond scholarship reconsidered: Toward an enlarged vision of scholarly work for faculty members. In K. J. Zahorski (Ed.), *Scholarship in the postmodern era: New venues, new values, new visions*. New Directions for Teaching and Learning, No. 90 (pp. 7–17). San Francisco: Jossey-Bass.

Schön, D. W. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books, Inc.

Shattuck, K. (2007). Quality Matters: Collaborative program planning at a state level. *Online Journal of Distance Learning Administration, 10(3)*. Retrieved from http://www.westga.edu/~distance/ojdla/fall103/shattuck103.htm

Suen, H. K., Logan, C. R., Neisworth, J. T., & Bagnato, S. (1995). Parent-professional congruence: Is it necessary? *Journal of Early Intervention*, 19(3), 243–252.

Zimmerman, W. A. (2010). *Quality Matters 2008–2010 rubric report*.

Zimmerman, W. A. (2013). *Analyses of 2011–2013 Quality Matters Rubric*.

End Notes

[1] O'Banion (1997) focus on learning as a key concept of the learner-centered movement has become intertwined with the scholarship of teaching.

[2] While this article focused on the continuous refinement of the QM Higher Education Rubric, QM provides other Rubrics and accompanying review processes:

- K–12 Secondary Rubric (Grades 6–12), which, after the first regular Rubric review, is now in the second edition
- Higher Education Publisher Rubric

• K–12 Publisher Rubric
• Continuing and Professional Education
Rubric