New Thinking, New Results
table of contents

2 Guest Commentary: Systems Thinking: Critical to Quality Improvement in Higher Education  
Julie Furst-Bowe

5 STEM: An Entrepreneurial Approach  
Keith T. Miller

8 Understanding Reliability in Higher Education Student Learning Outcomes Assessment  
Kenneth Royal

16 Using Active, Cooperative Quality Exercises to Enhance Learning  
James A. Griesemer
Dramatically shifting demographics, coupled with increased student expectations, continuous technological advances, and state and federal demands for increased completion rates are driving the need for significant improvement in higher education. It is extremely difficult to meet these challenges given the current environment of declining financial resources, and it is clear that institutions must reconsider traditional methods of operation and implement systematic methods for improving quality, efficiency, and effectiveness to remain viable in the global economy. Change must occur in all aspects of higher education, including academic programs, student support services, as well as administrative areas. Leaders must recognize that making changes in one aspect of an institution will have impacts on many other areas of the institution. For example, a decision to increase the international student population will impact several areas of the campus including English language programs, student housing, food service, as well as faculty and staff training and development programs. The University of Wisconsin-Stout (UW-Stout) has been involved in quality improvement for more than a decade, and senior leaders have realized that systems thinking is a major key to managing change and improving performance.

Systems thinking is a cohesive approach to management that views all key processes as parts of an overall system, rather than in isolation or as segments. Systems thinking is based on the idea that all key processes in an organization are interrelated. Understanding these relationships is critical to obtaining desired results, making targeted improvements, and achieving organizational effectiveness. When an organization is governed by systems thinking, work progresses at a faster, more efficient pace. Leaders with a systems-management approach guide synchronous actions across the entire organization, assuring alignment and integration of all units to maximize resources and productivity.

In a college or university setting, a systems perspective is essential for engaging the campus in setting goals, establishing priorities, allocating resources, identifying key performance indicators, and driving improvements. For example, if an institution sets a goal of increasing enrollment, all key processes and units, including marketing, recruitment, admissions, and financial aid must be aligned to achieve that goal. Resources must be deployed in these areas, as well as to the academic and student services units, to ensure adequate capacity to serve the increased number of students, both in and out of the classroom.

**Potential Challenges**

Implementing a systems perspective at a college or university, however, can be challenged by organizational structures, shared governance, faculty autonomy, and continued budget issues. Most higher education institutions continue to be organized in a traditional hierarchy, with several layers of management and numerous divisions and departments. The persistence of these functional silos, each with its own policies and processes, often leads to narrow vision, poor communication, and a lack of integration and alignment on campus-wide initiatives.

Although governance structures vary widely among higher education institutions, shared governance models that give faculty, staff, and students a voice in campus decision making are commonplace. Often, these internal stakeholder groups have very different motivations and priorities, making it difficult for institutions to move forward systematically with new initiatives or improvements to existing processes. In the United States, higher education is based on a tradition of academic freedom that allows faculty considerable autonomy in their teaching, research, and scholarly activities. This autonomy, however, can lead to pockets of faculty resistance and a lack of consistency when an institution is attempting to implement systematic methods for assessing student learning, using technology, or standardizing course evaluations across academic departments.

On a larger scale, budget limitations or funding formulas are often barriers to systems thinking and can dramatically affect how an institution establishes its priorities and allocates its resources. Frequently, budget cuts at public institutions include across-the-board reductions, employee furloughs,
or hiring freezes on vacant positions. Although these are some of the more manageable ways to deal with budget reductions, they are clearly not the most strategic and reflect a lack of systems thinking.

When across-the-board reductions are implemented, as opposed to strategic reallocations, institutions are unable to move forward with new initiatives. Overstaffed and understaffed departments and programs are treated equally, as are high-performing and low-performing employees. Priorities and resource alignment are compromised for the sake of convenience or fairness. These across-the-board actions conflict directly with systems thinking, which is based on strategic alignment, process management, and resource prioritization to drive continuous improvement.

**Baldrige and Systems Improvement**

Given all of these barriers, it is possible to develop and sustain a systems perspective and a culture of continuous and breakthrough improvement in higher education institutions. There are several models and frameworks that can assist campus leaders in developing this perspective and using systems thinking to benefit their institutions. The *Baldrige Education Criteria for Performance Excellence* provides a management model with a systems perspective for managing higher education institutions and their key processes to achieve results. The criteria also serve as the basis for the Malcolm Baldrige National Quality Award. First published in 1999, the education criteria have been used by postsecondary institutions across the United States for more than a decade. Most states and numerous other countries have established similar criteria and award programs based on the Baldrige criteria.

The education criteria are built on a set of interrelated core values and concepts, including visionary leadership, learning-centered education, and systems perspective. Within the Baldrige framework, a systems perspective is defined as the senior leadership focus on strategic directions and students. It means the senior leadership team monitors, responds to, and manages performance based on results, both short term and strategic. A systems perspective also includes using information and organizational knowledge to develop core strategies while linking these strategies with key processes and resources to improve both student and institutional performance.

**Baldrige Criteria in Use at UW-Stout**

UW-Stout began using the Baldrige criteria in 1999. In 2001, the school became the first higher education institution to receive the Malcolm Baldrige National Quality Award. One of 13 campuses that make up the University of Wisconsin System, UW-Stout enrolls approximately 8,800 in career focused undergraduate and graduate programs. The university continues to use the Baldrige criteria and was cited by the Academic Quality Improvement Program as a “national and international role model for quality in higher education.” Over the past decade, UW-Stout has demonstrated a systems perspective to performance excellence and has developed a culture of continuous improvement that has been tested by changing student demographics, declining state appropriations, and continuing turnover in key leadership positions. UW-Stout’s management approach has sustained its key performance results through changes in economic and market conditions. These performance goals are calibrated by best-practice benchmarks and competitive comparisons.

Although there are numerous components to UW-Stout’s quality management system, four components have been critical to the system and have been in place and refined continuously for more than a decade:

- An inclusive leadership system.
- A clearly defined set of student and stakeholder groups and understanding of their key requirements.
- A participatory planning process.
- An end-to-end system for measuring institutional performance.

UW-Stout’s inclusive leadership system was put into place in the mid-1990s, with goals of improving communication, trust, and decision making across the campus. The senior leadership team and its responsibilities were greatly expanded; the current senior leadership team has approximately 20 individuals, including administrators and representatives from faculty, staff, and student governance groups. This group meets every two weeks to review performance data, discuss issues, establish priorities, and serve as the key decision-making body for the campus. Members are responsible for communicating issues and actions with their representative groups. Members of the senior leadership team also serve on the strategic planning group for the campus. UW-Stout has implemented a comprehensive and robust strategic planning process beginning with a summer retreat attended by the senior leadership team and internal and external stakeholders, including alumni, community leaders, legislators, and employers. At this retreat, UW-Stout’s
mission, vision, and values are reviewed, performance is analyzed, emerging issues are discussed, and strategic priorities for the campus are drafted. Early in the fall, these draft priorities are shared with faculty, staff, and students at a series of listening sessions and through electronic communication.

Once the priorities have been finalized, action plans are created for each priority. Each action plan includes the responsible individuals or units, high-level steps, resources needed, a timeline and key performance indicators. Progress on action plans is monitored closely by the senior leadership team and there is a high level of accountability. Since this process was implemented, more than 60 action plans have been completed in areas such as globalization, e-scholar (laptop) deployment, applied research, and online program development. Action plans for university priorities are complemented by other university plans, including the academic plan, the integrated marketing plan, the affirmative action plan, and the IT plan.

Key Performance Indicators

Over the past decade, UW-Stout has refined its key performance indicators by focusing primarily on those that measure student engagement, progress, and success from the time students enter the university to after they have graduated and are employed in professional positions. Key student performance indicators include rates of applications, enrollments, retention, transfers, experiential learning participation, graduation, job placement, alumni satisfaction, and employer satisfaction with UW-Stout graduates.

These indicators were established through a comprehensive analysis of student and stakeholder requirements. They provide UW-Stout with a systematic view of the institution as students can be tracked at each stage of their college careers and beyond. Data from a number of sources, including surveys, are used to provide information for each performance indicator. These sources include the National Survey of Student Engagement or the ACT Alumni Outcomes Survey, and behavioral data, such as the amount of time students spend in campus laboratories or the percentage of students who participate in off-campus experiential learning programs. To assist in analysis, data are segmented according to student cohort, gender, race, or major program whenever appropriate. These types of analyses often help pinpoint specific problems in a program, process, or system.

Over time, UW-Stout has refined its use of comparative data and compares its performance to other U.S. and international institutions that have similar missions and programs. This is done to provide context for setting goals and analyzing institutional performance. When reviewing comparative data and stretch goals, the systems thinking perspective is critical to ensure that comprehensive strategies are formulated that consider all relevant factors involved in improving a specific performance indicator, such as student retention or graduation rates.

By using the Baldrige criteria and a systems-thinking perspective, UW-Stout has been able to demonstrate long-term progress in priority areas, such as increasing student enrollment, closing the achievement gap between majority and minority students, and increasing the number of students who participate in experiential learning programs. The campus has been able to achieve and maintain best-in-class status in areas that are key to its mission, including laboratory experiences, job placement rates, and employer satisfaction with graduates.

Systems thinking is based on the concept that all key processes in an organization are interrelated, and understanding these relationships is critical to obtaining desired results. The Baldrige criteria also require that senior leaders embrace systems thinking and promote that focus throughout the organization at all levels. The ultimate value in systems thinking in higher education is that it transcends institutional silos and provides campuses, such as UW-Stout, the ability to achieve institutional goals and sustain consistent performance improvement over time.

Editor’s note: This article is an update of Furst-Bowe’s June 2009 ASQ Higher Education Brief article, “Sustaining Performance Excellence in Higher Education.”

Julie Furst-Bowe

Julie Furst-Bowe is provost and vice chancellor of academic and student affairs at the University of Wisconsin-Stout (UW-Stout). Furst-Bowe co-authored Quality and Performance Excellence in Higher Education: Baldrige on Campus and Beyond Baldrige: What the First Institution of Higher Education to Receive the Malcolm Baldrige National Quality Award Has Learned in the Five Years Since. For more information about UW-Stout and its Baldrige award, visit www.uwstout.edu/mba/index.cfm.
The stage for the “2011 Advancing the STEM Agenda Conference” was set with these thoughts on how educators must join together to engage the workforce of the future in this vital area.

**STEM: An Entrepreneurial Approach**

Keith T. Miller

In a radio address, President Barack Obama said, “Today, more than ever before, science holds the key to our survival as a planet and our security and prosperity as a nation. It’s time we once again put science at the top of our agenda and work to restore America’s place as the world leader in science and technology.”

In many ways the president’s interest in science, technology, engineering, and mathematics (STEM) has energized the ongoing discussions among educators in K-12, higher education, and the workplace related to shaping tomorrow’s workforce. The fact that the world and its survival are becoming increasingly dependent on these interconnected fields has been recognized for quite a while, but reliable approaches for preparing students to succeed in that environment have not been adopted widely.

The challenge is to accept that there is no greater responsibility than to prepare the next generation of students to serve as the leaders of our future society. It is up to the current generation of leaders and educators to set priorities and to establish a new precedent. As advances in STEM fields will drive our society and economy, future generations that are well versed in STEM areas will have a disproportionate effect on the direction of initiatives and developments throughout society.

Preparing for this future will take perseverance, creativity, and an entrepreneurial attitude. One STEM student will be developed at a time, and success will be built from the experience of failures. Ultimately, growth in STEM areas through improved curricula and practical applications will affect every aspect of life.

**Getting Started**

Creating a strategic plan for our institutions that includes STEM topics is certainly a very important place to begin this journey. We must ask the question, “What will our society need in the future to optimize our potential as a society and as human beings?” The need for a
strategic approach becomes imperative when we begin to think about what knowledge and skills are necessary to address upcoming challenges. Furthermore, almost every unit in our organizations will need to be involved in shaping the plan, ensuring its unit goals and objectives fold into the overall scheme. Only a unified, high-level approach that ties directly to future STEM requirements can provide the platform necessary to move us to the improved performance necessary for success.

Setting the priorities is a great responsibility, and working with each other and the next generation of adults to implement plans related to these priorities is a privilege. This opportunity creates a need to think differently, which can set new precedents and affect the way we tackle the STEM agenda.

For instance, thinking differently might lead to the belief that any age can be the right time to start learning—even in the STEM fields. We shouldn’t exclude or dismiss any cohort group or any generation from STEM education. We need to demonstrate that learning truly is life long and incorporate appropriate approaches for all ages into our STEM plans.

Much research indicates that students should get involved in STEM courses in elementary school—or even as early as preschool and kindergarten. I would contend that STEM education parallels the study of languages where research findings support that greater success is attained when instruction begins at a young age. These statements probably reflect the ideal. Obviously, early exposure to STEM topics is best, but it is a relatively rare occurrence; however, we do have responsibility to make the ideal scenario happen to the greatest extent that we can.

**Fostering Interest and Learning**

It seems that a serious gap currently exists between the perceptions of Americans toward STEM topics and the need for a greater focus on them to achieve success in the future. Many students in the United States do not seem to enjoy math or science. The challenges educators will need to overcome to prepare the workforce are immense.

This underlying disconnect appears related to what motivates our willingness to tackle education. Apparently, we are more likely to spend time studying subjects we enjoy and think are fun instead of those that will benefit society and generate high-potential career paths for us.

Some of this lack of enjoyment, no doubt, is associated with the effort required to master math and science, which generally are viewed as more difficult subjects. Children and adults often experience failure when learning and applying science and math, and they can become discouraged. They focus more attention on topics where they feel more capable and successful. The innate tendency to avoid failure creates a barrier to learning that may be difficult to overcome unless we adopt new strategies that foster continuing effort.

The key is to recognize the value of failure as a process that provides a foundation for future learning. Furthermore, successful strategies ensure that excessive failure is not acceptable; a certain amount of failure makes us better, but too much failure undermines our confidence and brings out our self-protective behaviors. No one likes hitting his/her head against the wall continually, but many people enjoy beating the odds by tackling and overcoming an obstacle. To encourage engagement in STEM education, we must find the fine line between having too much failure and just enough.

As educators, we should attempt to speed up the mistakes because mistakes are going to happen. Temporary depression will occur in conjunction with these failures; it happens to all of us. How we respond as teachers, parents, and mentors is what makes the difference and matters most. We need to remember, “success is 99% failure” and be prepared to coach students through the down times to success.

Our goal should not be to turn every student into a mathematician, scientist, or an engineer. Instead, we should focus on adding one math or science student at a time to our institutions. We can achieve this objective if we implement an innovative approach based on the entrepreneurial model.

**Implementing the Entrepreneurial Model**

From a business perspective, much of the growth, when there is growth in the United States economy, is due to the work of entrepreneurs and small business owners. Successful entrepreneurs are able to find a niche and take advantage of it. Many entrepreneurs have failed numerous times
before achieving success. If we think differently about education, we can establish a K-16 framework based on this same model that will drive improved workforce STEM preparation. Certainly, STEM courses are very important at every level of education.

We need to create a culture that presumes young children can enjoy and succeed when learning mathematics and other STEM subjects. This culture must be built on the belief that we have the approaches in place to stimulate interest and generate capability in these areas.

One way to capture the attention of students and generate dedicated effort to pursue a degree in the STEM field is to clarify the benefits of careers in these fields. Demonstrating the relatively higher earnings potential, associated better quality of life, and increased job security can be a strong incentive. Individuals with STEM skills entering industry tend to have higher salaries than those in other disciplines. Often those with two-year STEM degrees will make more entering the workforce than those with other types of four-year degrees. Furthermore, I anticipate that STEM graduates are likely to assume greater leadership roles as organizations around the world become more flat and expect middle managers to possess more technical skill.

What can we do to make STEM courses more attractive to students? How can we increase student success in these traditionally challenging courses? At the college level, we can integrate STEM courses with case studies and experiential activities to improve their effectiveness. We also can incorporate some aspects of STEM across the curriculum in English, social studies, and geography. For example, courses might include: “Psychology of Science and Technology,” “Politics of Science and Technology,” “Economics of Science and Technology,” “Business of Science and Technology,” and “Sociology of Science and Technology.”

Our efforts to engage students need to begin long before college, however. Some wonderful initiatives are available across the United States with an influx of summer camps, school partnerships, and mentoring programs with professionals, as well as public school/private industry partnerships. There are Saturday programs, speaker series, company tours, scholarship programs, grants, and more.

**Summary**

Was the success of Albert Einstein, Steve Jobs, or Bill Gates a straight line of success? Probably not, there was likely a bump or two (some failure) along the way. If we think about STEM across the curriculum as an entrepreneurial venture and revisit how to educate holistically, we can challenge ourselves to meet the STEM needs around the world by the year 2025.

Dr. Keith Terrence Miller is the 13th president of Virginia State University. He worked in corporate America including stints at Pennsylvania Power and Light, Proctor & Gamble, and the Omnistaff Corporation. His professional academic career includes various instructional and administrative roles at Fairleigh Dickinson University, Quinnipiac College, Niagara University, and the University of Wisconsin, Oshkosh. In 2004, he was named president of Lock Haven University of Pennsylvania and later was named president emeritus. He can be contacted at ktmiller@vsu.edu.
Abstract
The accountability movement in higher education has required virtually all faculty and most administrators to assess student learning outcomes (SLO). Unfortunately, not everyone who is required to conduct assessments is an expert in social science research methods or educational assessment. This is particularly problematic because fundamental concepts often become distorted and the misuse of such concepts can have a very negative impact on both local assessment findings and even the larger field itself. One fundamental concept that is not particularly well understood, but ought to be, is reliability.

This article addresses some of the most common misconceptions about reliability in the higher education SLO arena and encourages practitioners to be more attentive to the details of calculating, interpreting, and reporting estimates of reliability. The article provides an overview of reliability and the factors that influence it, discusses the most commonly used types of reliability in SLO assessment, and provides guidance on how to effectively interpret and report such measures. The information presented should be particularly helpful to SLO practitioners who need a brief primer on reliability, ponder how to construct better assessment instruments, have an interest in...
making appropriate inferences about assessment results, and are concerned with reporting findings accurately and responsibly.

Introduction

Most faculty members and administrators involved in the higher education enterprise are expected to participate in SLO assessment. This is true for faculty from virtually every academic discipline, as well as administrators in various units within a college or university. Unfortunately, whenever universal requirements are in place there is often a great deal of confusion as to best practices, especially for those who are not necessarily experts in educational assessment or social science research methods. One such topic where there is a good bit of confusion in the SLO assessment arena is the notion of reliability. Most know that high reliability estimates are desirable, but many often assume assessments, which do not have high levels of reliability, are of poor quality. This is not necessarily true, as there are multiple factors that affect the reliability of any assessment. In fact, the lack of understanding about reliability is so pervasive that many fine instruments are regularly dismissed and valuable time and energy are wasted trying to revise them so that they yield more reliable measures. With a proper understanding of reliability, such revisions may not be necessary.

This article addresses some of the most common misconceptions about reliability and encourages those who work in higher education SLO assessment to refrain from hastily dismissing results and instruments without careful consideration. Additionally, this article discusses the most prevalent types of reliability used in SLO assessment, how to appropriately interpret these estimates, and report these estimates responsibly to various audiences.

Overview of Reliability

Generally speaking, reliability refers to the extent to which scores/results are repeatable and stable, that is, produce the same results on repeated trials. It is important to understand that any measurement will always contain some amount of error. The extent to which the error is minimized determines the reproducibility and stability of any set of scores. That is, when little error is present scores are deemed to have high reliability; likewise, when a great deal of error is present, scores are deemed to have low reliability.

Before proceeding further, let us briefly discuss reliability in our daily lives. Consider, for example, a piece of equipment or machinery that continues to work without any problems. This piece of equipment is said to be reliable. Likewise, a person may be considered a reliable witness if he or she provides a consistent testimony after cross-examination during a legal proceeding. An athlete may be considered reliable if he or she typically performs in a consistent manner during a sporting competition. The world around us is full of examples that assume a uniform definition of reliability. With that said, it is no wonder a significant number of assessment practitioners carry over such conceptualizations into the SLO assessment arena by asking such questions as “How reliable is your instrument?” or “How can I make my instrument more reliable?” It should be made abundantly clear that psychological instruments used to measure latent traits (e.g., ability, knowledge, attitudes, etc.) do not possess the property of reliability. Therefore, there is no such thing as a reliable instrument. Only the results produced from an instrument have the property of reliability.

With results possessing the property of reliability, and not the instrument, it is easy to understand how multiple factors could impact reliability. In fact, there are three key factors that determine the reliability of any quantitative assessment:

- The characteristics of the instrument.
- The conditions of administration.
- The characteristics of the sample.

It is the interaction among these three factors that determine how reliable the results are from any assessment. Let us investigate each of these three factors separately.

Characteristics of the Instrument

The first factor is the characteristics of the instrument. These characteristics include the length of the instrument, types of items employed, and quality of the items. As a rule of thumb, more items tend to produce more reliable scores than fewer items. The types of items employed are also important. Objective items, such as a multiple-choice
test or a survey with a limited number of viable options, are much more likely to produce reliable scores than subjective items such as essays. This is largely for two reasons. First, objective items reduce scorer inconsistency as a source of measurement error. Second, more content can be covered, which reduces the chance that students will luckily or unluckily receive a particular writing topic. Generally speaking, an instrument that contains a larger number of objective items will produce more reliable scores than an instrument that contains fewer and more subjective items.

Item quality is also paramount. Items should be clear, unambiguous, and worded in an unbiased manner. Regardless of whether it is a test or a survey, items should discriminate well. Discrimination refers to the extent to which items are able to discern differences among examinees or survey participants. For example, consider a test that contains nearly all easy items. We would anticipate that even the less able students would answer many items correctly, thus resulting in little information regarding what test takers truly know (or can do). By contrast, a test that contains a sufficient mix of easy, moderate, and difficult items is much more likely to distinguish among those who are more able (or knowledgeable) from those who are not. The same is true for a survey. When we attempt to measure a latent trait such as an attitude, perception, belief, etc., we need items that can better draw out how much of the latent trait an individual possesses. This notion is particularly problematic for many working in SLO assessment, as many instruments provide questions that are very easy for respondents to endorse. Most student satisfaction surveys are notorious for this. The problem is that it is difficult to make reliable distinctions among the results when students tend to agree with every item appearing on the instrument. It is therefore imperative that items vary sufficiently in difficulty.

**Characteristics of the Sample**

The third factor is the characteristics of the sample. This factor is perhaps the most misunderstood of the three factors that influence reliability. We often assume that any good sample will do. This assumption tends to have real consequences for assessors. In much the same way that a good instrument should have a wide array of items that vary in difficulty, we need samples with respondents that also vary in whatever characteristics we are attempting to measure (such as ability, attitudes, etc.). For example, suppose a well-targeted test was administered to two groups of college seniors. In the first sample, the range of ability across the students was quite large. That is, there were clearly some students who were far less/more able than the others. In the second sample, the range of ability was quite restricted, as virtually every student in the class was of comparable ability. Although the instrument is sound and the conditions for administration were identical, the ability among those in the samples is quite different. This means the scores produced from the first sample (greater range of ability) will likely yield higher reliability estimates than those from the second group. Bear in mind that reliability investigates the extent to which the test is capable of making reliable distinctions among the persons in the sample with respect to the latent trait (in this
case, ability) that is measured. Stated another way, the smaller the variance in scores, the less reliable the scores. Conversely, the larger the variance in scores, the more reliable the scores.

These three factors impact virtually all quantitative assessments. It is possible therefore, that an instrument with sound psychometric properties can still produce poor measures of reliability. It is also possible to administer an assessment in exactly the same manner under the exact same conditions and still have poor estimates of reliability. Finally, it is also possible that one can administer an instrument to a heterogeneous sample and still have low reliability. Separately, each of these factors is very important, but as stand-alone indicators of quality they can be deceiving. A problem with any one of these factors can yield poor reliability estimates. When we are presented with reliability estimates that are less than desirable, we need to consider each of these factors and determine how we can make improvements. That is, do we need to revisit our items, do we need to go to greater lengths to ensure a cleaner administration of the assessment, or do we need to reconsider the variance in ability (or other latent trait) in our sample? These are the types of questions that can go a long way in improving the reliability of any results. Next, let us turn our attention to types of reliability.

Types of Reliability

As mentioned previously, calculating reliability is about understanding and accounting for different sources of error. Various types of reliability calculations are available to estimate a particular type of error, but the choice as to which type of estimate to use depends upon one’s purpose. Those who work in the SLO assessment arena often find themselves concerned with three types of reliability in particular: stability, internal consistency, and agreement.

Stability

Stability refers to the extent to which scores are reproducible as a result of administering the same instrument to the same sample of participants at different points. A relevant example to SLO professionals might include a math instructor who administers a survey at the beginning of the semester that measures students’ social values. The instructor then administers the same instrument to the same group of students at the end of the semester. As the math course did not address issues regarding social values, it would be reasonable to assume that students’ social values did not change much over the duration of the semester. If the results were highly correlated (.6 or higher), one would assert that test-retest reliability is high.

Another type of reliability that measures stability is alternate-forms reliability. This type of reliability involves using two comparable instruments that are intended to measure the same construct to determine the extent to which scores between the two instruments correlate. A classic example of alternate-forms reliability is when an instructor administers multiple versions of a test. If the tests are truly psychometrically comparable (e.g., same content, comparable difficulty, same types of scales), the scores produced from each instrument should be highly related. Extremely high levels of reliability would suggest that it makes no difference which version of the test a student receives, as he/she would not be unduly advantaged or disadvantaged either way. Alternate-forms reliability is also helpful for SLO professionals who are interested in instrument development. For example, suppose one wishes to create an item bank of survey items. Instead of administering the same limited pool of items each time, assessment professionals could use alternate or perhaps mixed forms of the instrument. If equated procedures are sound, assessment professionals could generate a number of comparable instruments with minimal error. Although a great deal of promise exists for these techniques in the SLO arena, to date this avenue of psychometrics has rarely been explored in SLO research and practice.

Internal Consistency

Internal consistency refers to the extent to which the items that comprise an instrument are correlated. This type of reliability is perhaps the most commonly referenced in SLO assessment. Suppose, for example, that students are given an attitudinal survey about campus climate. If students tend to agree with items such as “the campus provides a positive environment…” and disagree with items such as “the campus provides a negative environment…” then the results would be highly correlated, as they tend to tell the same
story. This is what is meant by internally consistent. It should be noted, however, that there are several ways to examine internal consistency.

One method for determining internal consistency is to divide the instrument in half and correlate the two sets of items. This type of estimation is called split-half reliability. In instances where there is a correct and incorrect answer (such as a test), the Kuder-Richardson method is helpful. Unlike the split-half method where the researcher must determine where to divide the test, the Kuder-Richardson method overcomes this problem by calculating an average correlation based on all possible split-half estimates.

Another tool that may be helpful for those in SLO assessment is the Spearman-Brown Prophecy Formula, which calculates the impact on reliability from adding more items to an instrument. In general, increasing the number of items will improve reliability as long as the new items are consistent with the original items. The formula predicts that reliability increases more slowly than the increase in the number of items. This can be particularly valuable for those creating survey instruments who are concerned with survey length, since a great deal of research suggests that too many survey items will result in survey fatigue for participants, thus increasing the likelihood of missing data.

Coefficient alpha, often referred to as Cronbach’s alpha, is another method used for testing internal consistency. Cronbach’s alpha estimates internal consistency in much the same way as the previously noted Kuder-Richardson method. The primary difference is Cronbach’s alpha produces estimates of internal consistency when the response choices contain a range of options, such as a Likert scale (e.g., strongly disagree to strongly agree).

Agreement

Agreement refers to the extent to which agreement exists among ratings. There are several methods for which a coefficient of agreement can be determined, most of which involve either a correlation or a percentage of agreement. In the SLO assessment arena, it is common to recruit judges to provide ratings of a performance. For example, a panel of reviewers may use a rubric to judge the merits of writing samples based on certain criteria. In such cases, it would be helpful to know the extent to which the reviewers are consistent in their ratings, as reviewers who are unduly harsh or lenient could bias the quality of the ratings. This is referred to as interrater reliability. Some researchers opt to simply provide descriptive counts and percentages of common responses. Some researchers will use the statistic kappa to report levels of interrater reliability when there are two raters. Unfortunately, the kappa statistic assumes no ordering among the responses and treats the responses as nominal. For this reason, SLO assessors need to be aware of this approach as it is a statistical violation when used for this purpose.

Others may prefer to use Pearson’s correlation because they assume the rating scale is interval in nature. A long history of measurement research has shown agreement scales are indeed ordinal in nature; therefore Pearson’s correlation would be an inappropriate statistic for ordinal scales. Based on level of measurement, it seems Spearman’s rho would be the most appropriate correlation. Unfortunately, this technique has its share of problems when used for agreement correlations as well, because this method only correlates the consistency of the pattern of responses. This problem is highlighted in the following example: Two reviewers provide the following ratings to five criteria:

Reviewer one: 1, 3, 1, 3, 1
Reviewer two: 2, 4, 2, 4, 2

Spearman’s rho would provide a reliability estimate of 1.00, indicating perfect reliability. This is deceiving however, as the two reviewers failed to actually agree on anything. Instead, it is merely the pattern of agreement that was comparable across the two raters. Testing for absolute agreement using intra-class correlations would be a more appropriate method because it takes into consideration both the pattern of agreement and the matching of identical responses. When taking absolute agreement into account, the intra-class correlation of these scores is .29, revealing a very different finding than Spearman’s rho, which suggested perfect agreement. Without careful attention to detail and truly understanding what is taking place with reliability estimation, SLO assessment professionals may produce deceptive findings inadvertently. This could have very
negative consequences, as untruthful results may be used as the basis for decision making, thus resulting in a different set of problems.

Of course, other more advanced techniques are available as well. More recently, methods such as Generalizability Theory\textsuperscript{16} and the many-faceted Rasch model (MFRM),\textsuperscript{17} a form of item response theory, have become widely used in educational measurement literature. Most measurement experts contend that both of these techniques are superior to traditional statistical approaches (called classical test theory) for analyzing such performance assessment data.\textsuperscript{18} Although these particular methods are beyond the scope of this paper, it is important that assessment professionals begin to investigate more modern methods with stronger theoretical underpinnings for analyzing data.

Interpreting Reliability

Before interpreting a reliability estimate, it’s important to clearly understand that no measurement is perfect. Even widely-held “concrete” facts that we have come to know about assessment are sometimes questionable. For example, consider a test where a student answered 40 out of 50 items correctly. It is true that the raw score count of 40 correct items does not contain any error. When we say something more about this, such as making an inference about someone’s ability, then all of a sudden error is introduced. How can we be sure that someone else of the same ability would also mark the same 40 items correctly? Can we be sure that this same student would have provided the same answers if we gave him or her the test again? What about the circumstances surrounding the test? What if the student was extremely tired from studying most of the previous night and is not functioning at full capacity at test time? We can provide countless examples and scenarios, but the point is that any inference we make about a score is just that, an inference. It is susceptible to error. It should be clear that error exists in all measures; however, it is our job to control it as best we can.

What exactly does a reliability estimate tell us? If an assessment professional analyzes his or her survey data and finds Cronbach’s alpha is estimated at .85, this means he or she can say that an estimated 85% of the observed variance in ratings is due to systematic differences in participant responses, with 15% due to chance differences or, alternately, that 15% of the observed variance is due to measurement error. The same interpretation could be made with a reliability estimate of test scores or ratings from a rubric as well. Once we are able to interpret the meaning and magnitude of any reliability estimate, it is helpful to then understand whether or not (and to what extent) the estimate is considered acceptable.

Obtaining Sufficiently Reliable Results

What are sufficiently reliable results? This answer will vary depending upon whom one asks as well as the purposes (and methods) of the assessment. In testing scenarios, how reliable the scores are depends mainly on what is at stake. According to Nunnally and Bernstein,\textsuperscript{19} a reliability estimate of 0.90 should be the minimum when an exam is high stakes in nature. Reliability estimates above 0.80 are considered reasonably reliable, and anything less than 0.80 would be a poor foundation for drawing significant conclusions and decision making. Others in the high-stakes testing arena might have more liberal opinions. For instance, de Klerk\textsuperscript{20} suggests estimates above 0.80 are exceptionally good, and anything above 0.70 is acceptable. Although there is room for some disagreement about what is acceptable, nearly all measurement experts will agree that anything less than 0.60 is unacceptable in the high-stakes arena.

Of course, when an exam is not high stakes, an estimate of 0.70 or higher may more than suffice, while an estimate above 0.60 would be considered respectable. In survey scenarios, any estimate above 0.70 is generally considered quite reasonable for most audiences;\textsuperscript{21} however, individual preferences may lead one to expect slightly higher (0.80) or lower estimates (0.60). George and Mallery\textsuperscript{22} have provided a commonly accepted rule of thumb for Cronbach’s alpha that classifies reliability results, as shown in Table 1. Of course, much of this depends on the specific nature of the assessment as well.

Rubrics, however, may be a bit of an exception. As mentioned previously, numerous authors have published works boasting of reliability estimates meeting or exceeding 0.90 when utilizing a rubric. For the most part, this is misleading (as in the case mentioned previously with Spearman’s rho used to analyze data obtained from a rubric).
Consider the three factors mentioned earlier in this article. Rubrics contain far fewer criteria to measure, namely fewer items, than a traditional test or survey. Rubrics are also subjective in nature, regardless of the means taken to make it more objective. Even the conditions for which the raters make their judgments are susceptible to error, as rater drift is a very real phenomenon.23,24 Given these problems, it is not feasible or responsible to contend that agreement estimates should be 0.90 or higher. In fact, estimates greater than 0.60 might be exceptionally high25 and estimates between 0.41 and 0.59 might be acceptable if using intra-class correlations, for example.

Whenever reliability results are judged as less than acceptable, there is a need to revisit the three factors impacting reliability: the instrument, conditions of administration, and characteristics of the people who took or will take the instrument.

Table 1: Interpretation of Reliability Based on Cronbach’s Alpha

<table>
<thead>
<tr>
<th>Cronbach’s Alpha</th>
<th>Internal Consistency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9 ≤ α</td>
<td>Excellent</td>
</tr>
<tr>
<td>0.8 ≤ α &lt; 0.9</td>
<td>Good</td>
</tr>
<tr>
<td>0.7 ≤ α &lt; 0.8</td>
<td>Acceptable</td>
</tr>
<tr>
<td>0.6 ≤ α &lt; 0.7</td>
<td>Questionable</td>
</tr>
<tr>
<td>0.5 ≤ α &lt; 0.6</td>
<td>Poor</td>
</tr>
<tr>
<td>α &lt; 0.5</td>
<td>Unacceptable</td>
</tr>
</tbody>
</table>

Conclusions

Reliability is one of the most commonly reported, yet misunderstood notions in the higher education SLO arena. This article introduced the concepts of reliability and the factors that influence it, the most commonly used types of reliability in SLO assessment, and how to interpret and report such measures effectively. The information should be particularly helpful to SLO practitioners who need a brief primer on reliability, ponder how to construct better assessment instruments, have an interest in making appropriate inferences about assessment results, and are concerned with reporting findings accurately.

References
2. Bruce Thompson, Score Reliability; Contemporary Thinking on Reliability Issues, Sage, 2003.

Dr. Kenneth Royal is a psychometrician at the American Board of Family Medicine and an adjunct professor at the University of Kentucky. He has published and presented more than 90 papers spanning multiple disciplines such as higher education, psychometrics, medicine, psychology, etc. His primary research interests include the application of various quantitative methods for improving quality. He can be contacted at kdroya2@uky.edu.
Using Active, Cooperative Quality Exercises to Enhance Learning

James A. Griesemer

Abstract
Active, cooperative quality learning exercises were used to enhance learning in an undergraduate operations/supply chain management course. The quality tools and techniques supported course topics while also adding to students’ first-hand knowledge of quality management. Such exercises were found to improve students’ critical thinking and problem-solving skills significantly. Based on this early success, additional active, cooperative quality learning exercises are being developed for use in other business courses.

Introduction
Few small to medium-sized colleges offer courses in quality. This means students are only exposed to quality as a topic in an operations/supply chain management course if it is required or offered. In a popular textbook, F. Robert Jacobs and Richard B. Chase’s Operations and Supply Chain Management, the chapter on six-sigma quality with its appendix on process capability is one of 20 chapters. In a traditional semester if this chapter is included in the course it means quality is only formally taught for one or two class meetings.

Good teaching has been defined as instruction that leads to effective learning, which, in turn, means thorough and lasting acquisition of the knowledge, skills, and values the instructor or the institution has set out to impart. Unfortunately, most students cannot stay focused throughout a lecture. After about 10 minutes their attention begins to decline, first for brief moments and then for longer periods, and by the end of the lecture they are taking in very little and retaining less. A classroom research study found that immediately after a lecture, students recalled 70% of
the information presented in the first 10 minutes and only 20% of that from the last 10 minutes.3

It has been found that students’ attention can be maintained throughout a class session by giving them something to do periodically. Though many different activities can be used, the most common is the small group exercise. Such active, cooperative learning exercises may address a variety of objectives. These include problem solving, as well as analytical, critical, and creative thinking as long as the exercises are conducted to foster positive group member interdependence, individual member accountability, face-to-face interactively, use of teamwork skills, and regular team self-assessment.4

Approach

Incorporating active, cooperative quality learning exercises in a course requires instructors to modify their teaching strategy in a number of important ways but the most critical is their roles as educator, mentor, and facilitator. Research confirms the effectiveness of active, cooperative learning. Compared to students taught with conventional methods, cooperatively taught students tend to exhibit better grades as well as better analytical, creative, and critical thinking skills among other traits.3

The pedagogical sequence used for employing active, cooperative quality learning exercises for enhancing learning included these four key steps:

1. Planning: Setting desired course outcome goals in terms of knowledge gained, comprehension of terms and concepts, ability to use knowledge and analyze results, and linking them to various quality concepts, tools, and techniques.

2. Coaching: Supplying students with supplementary instruction on quality tools, techniques, and resources and providing one-on-one mentoring on how to use them.

3. Evaluating: Grading the students’ work using a rubric that is agreed upon by the students. The rubric emphasized not only quality knowledge but also the use of various quality tools and techniques as well as the topic under study at the time.

4. Assessing: Collecting students’ feedback on their experiences and recommendations for improving both the active, cooperative quality learning exercises and the course in general.

The operations/supply chain management course is an introductory course covering numerous topics thus resulting in it having a wide spectrum of outcome goals. The quality tools and techniques used were selected to support the topics studied while at the same time adding to students’ learning of quality management. For example, the first exercise is a variation of Dr. Deming’s red bead experiment.5 Students begin by taking random samples from a large box containing eight different colored plastic beads. The first sample has a size of 10, the second has 25, and the third contains 100 beads. The students then construct a histogram for each sample. This helps illustrate and emphasize the importance of proper sample size. Then they continue with a scaled-down version of the red bead experiment with the students alternating between serving as the foreman and the worker.

The second exercise focuses on the challenge of controlling the variation of the weight of coins in ancient Greece. Students learn the difference between variable versus attribute data and the use of control charts. The remaining eight exercises are brief, self-contained case studies written by the author and are based on former students’ internship experiences at various local businesses. They require students to use a variety of quality tools such as check sheets, cause and effect diagrams, Pareto charts, matrix diagrams, and radar charts, and then make recommendations to management. All of the exercises were designed to take 25 minutes or less to complete and the majority take approximately 10 minutes. The second quality case study is featured in the sidebar, “Case Studies Provide Experience With Quality Tools.”

Student Experience

During the first course meeting students were randomly assigned into two- or perhaps three-person groups by the instructor. One member was designated as the recorder and each was told he/she might be called upon to present findings in the form of an executive summary. A sample executive summary was provided to the teams and they were mentored closely in writing the first several summaries. When the exercises would occur was not announced, and the first exercise was conducted in the second week of the course with the remaining nine exercises occurring one per week over the next 12 weeks.
Initially some students resisted being assigned to partner. They claimed it forced them to work with students they did not know. Others viewed the exercises as just more homework. Within a short time though, students viewed the exercises as an interesting break from the usual routine of lecture followed by homework help. By the end of the course, students’ creative thinking and problem-solving skills had improved significantly. They also developed working knowledge of several traditional quality tools and, more importantly, an appreciation for quality management and especially the Deming Cycle focusing both on short-term continuous improvement and long-term organizational learning. Student feedback included numerous comments identifying

**Case Studies Provide Experience With Quality Tools**

Here is one of the eight case studies used as an active, cooperative learning exercise in the undergraduate operations/supply management course at Mount Saint Mary College.

Corey is a recent college graduate, who has been hired by a local, specialty candy company that makes a considerable amount of their profits from manufacturing and selling multi-colored, sugar-coated chocolate candy. The general manager feels the control of the mixture of the different colored candies that go into the various sized packages is not right. Customers have been complaining that they have not been receiving candies of all colors. As a result of these complaints, he has asked Corey to look into it. After talking with his manager, Corey decides he has to collect some data on the number of different colored candies in one specific package, in this case, the 14 oz. family bag. His manager advises him that data is the raw material of quality control and he has to be especially conscientious to make sure the data he collects is good.

Before he collects any data, Corey needs to understand how the candies are put into the packages. He talks with Samantha, who is the assistant production manager, and she tells him there are separate candy-making machines for each color and they all discharge into a large mixing bowl that, in turn, feeds the packaging machines. There are multiple manufacturing lines for some colors and the machines run at different speeds because colors dry at different rates. There are no standards for how many of each color candy there should be in a package. Samantha thinks this would be a good idea because otherwise only well running and fast drying color candy will make it into packages. Corey spends the next hour or so watching the manufacturing process and testing the candies.

Corey reports back to his boss and mentions Samantha’s idea about introducing standards for the number of candies of each color in each package. His boss likes the idea and suggests Corey look into using control charts to help him establish the standards. Later, Corey calls the quality assurance group and they offer to help him. The first thing they suggest is that Samantha set up all of the candy manufacturing machines and then let them run without interference for a while. Next, they tell Corey he must determine if he is working with variable or attribute data. After a while they tell Corey to find a place somewhere between the mixing bowl and the packaging machines where he can collect samples without disturbing the process. Since he is working with the 14 oz. package they tell him to first randomly collect 100 test specimens per sample and five samples over the next hour for use in making a histogram.

Answer the following questions:

- Is Corey working with variable or attribute data? Explain briefly why.
- Construct the histogram. What does it tell Corey and why is it important to know?
- Select a control chart to use. What factors influenced your decision? Determine the control limits.
- Prepare an executive summary for the general manager.
the exercises as the most interesting and valuable part of the course.

Faculty Experience

Incorporating active, cooperative quality learning exercises into a course requires instructors to modify their teaching strategies. Their role changes from serving primarily as a presenter of knowledge to also being a facilitator and mentor. Active, cooperative learning is not easy to implement and instructors who simply assign students to work in teams do not realize its full benefits. Team exercises that do not ensure individual accountability can result in some students getting credit for work done by their more industrious and responsible partners.

Introducing active, cooperative quality learning exercises into a course is a challenging undertaking. Lessons learned from this recent attempt include:

- Design active, cooperative quality learning exercises so the learning objectives complement the learning objectives of the course topic. The first exercise based on Deming's red bead experiment teaches students not only important lessons about variation but also management.

- The exercises must include challenging assignments for all team members, have a reasonable required completion time, and specific and measurable outcomes. When properly designed, they can create alternative ways for covering topics and even serve as a means for introducing additional topics.

- It is especially important that instructors design time for active, cooperative quality learning exercises into their course syllabi. Using students to test exercises before they are used in a course helps identify any potential problems early.

- Select student teams that are heterogeneous in ability to encourage student involvement and interaction. Limit teams to two or possibly three students to ensure no student is left out of the group process.

- Ensure that active, cooperative quality learning exercises are more challenging than individual assignments. The level of challenge should not be increased by making the exercises longer but by requiring higher-level thinking skills thus contributing to the growth of students' critical thinking skills.

- Be ready to teach students to how to work effectively in teams. One successful approach is to collect anonymous comments from students after the first exercise about any problems they experienced and then brainstorm possible solutions in class before the start of the second exercise.

- Hold students accountable in as many ways as possible. This could include having the instructor select, without notice, which student will report the team’s findings and asking team members to grade each other’s contribution with the average being the team’s grade.

These findings support the premise that adding active, cooperative quality learning exercises to an undergraduate course in operations/supply chain management course has many benefits. These exercises can help enhance students’ problem-solving skills and creative thinking. They can also be used to complement course subject topics, thus enriching students’ course experiences.

The benefits were also found to extend beyond the course. Faculty teaching the required business capstone course reported that students used quality tools to analyze management case studies and then reported their findings in executive summaries. The faculty felt the use of quality tools enhanced the students’ learning experience and better prepared them for entering the workforce.

Based on this early and encouraging success, plans are now underway to develop additional active, cooperative quality learning exercises for use in other business courses. Initial feedback from students indicates they especially enjoy the hands-on experience of active, cooperative quality learning exercises, which help them retain information longer.

Conclusions

Both instructors and students reported numerous benefits of incorporating active, cooperative learning quality exercises into an undergraduate
operations/supply chain management course. It required an extensive change in course design and made the instructor assume a broader role in the classroom, including that of a facilitator and mentor. Students viewed the exercises as an enjoyable break in the usual class routine and appreciated learning about various quality problem-solving tools and techniques.

References

James A. Griesemer is an associate professor of business at Mount Saint Mary College located in Newburgh, New York. Griesemer teaches courses in quality assurance, production systems, operations management, and management science. His research interests include the use of mathematical models and software applications to solve complex business-related problems. Prior to becoming a professor he worked in research and development for International Paper Company. Griesemer holds a doctorate in management from Pace University, an MBA in financial management from Long Island University, an MS in materials science from Fairleigh Dickinson University, and a BS in engineering from SUNY College of Environmental Science and Forestry. Contact him via email at James.Griesemer@msmc.edu.
Call for Articles

Quality Approaches in Higher Education

The American Society for Quality's Education Division has launched a new bi-annual, online, peer-reviewed journal called Quality Approaches in Higher Education. The editorial review team actively encourages authors to submit papers for upcoming issues.

The purpose of this publication is to engage the higher education community and the ASQ Education Division membership in a discussion on topics related to improving quality in higher education and identifying best practices in higher education and to expand the literature specific to quality in higher education topics. Quality Approaches in Higher Education welcomes faculty from two- and four-year institutions, including engineering colleges, business schools, and schools of education, to consider submitting articles for review.

The following types of articles fit the purview of Quality Approaches in Higher Education:

- Case studies on how to improve quality in a college or university.
- Conceptual articles discussing theories, models, and/or best practices related to quality in colleges and universities.
- Research articles reporting on survey findings such as a national survey on students’ attitudes toward confidence, success in college, social networking, student engagement, access and affordability, etc.
- Case studies or conceptual articles providing institutional perspective on process development and maintenance methodology at colleges or universities.
- Case studies or conceptual articles addressing issues such as the role of faculty and administrators in quality systems.
- Case studies, research studies, or conceptual articles focusing on accreditation issues.
- Case studies demonstrating best practices using the Baldrige Education Criteria for Performance Excellence, including experience and recommendations for successful implementation.
- Case studies, research studies, or conceptual articles on scholarship of teaching, enhancing student learning, learning outcomes assessment, student retention, best practices for using technology in the college classroom, etc.

In particular, we are looking for articles on the following topics: using assessments for continuous improvement and accreditation, showing how use of the Baldrige framework can increase student success, increasing engagement and quality of learning through lecture capture and other technologies, dealing with rising costs without jeopardizing learning, sponsoring programs for helping graduates gain employment, and merging research with practice (action inquiry).

Articles generally should contain between 2,500 and 3,000 words and can include up to four charts, tables, diagrams, illustrations, or photos of high resolution. For details, please check the “Author Guidelines” at http://www.asq.org/edu/2009/09/best-practices/author-guidelines.pdf.

Please send your submissions to Dr. Cindy Veenstra at qahe@asqedu.org.
Author Guidelines

Quality Approaches in Higher Education

Quality Approaches in Higher Education is peer reviewed and published online by the Education Division of the American Society for Quality (ASQ). The purpose of this publication is to engage the higher education community and the ASQ Education Division membership in a discussion of topics related to improving quality and identifying best practices in higher education and to expand the literature specific to quality in higher education topics. We will consider articles that have not been published previously and currently are not under consideration for publication elsewhere.

General Information

Articles in Quality Approaches in Higher Education generally should contain between 2,500 and 3,000 words and can include up to four charts, tables, diagrams, or other illustrations. Photos also are welcome, but they must be high resolution and in the format described later in the “Submission Format” section.

The following types of articles fit the purview of Quality Approaches in Higher Education:

- Case studies on how to improve quality in a college or university.
- Conceptual articles discussing theories, models, and/or best practices related to quality in colleges and universities.
- Research articles reporting on survey findings such as a national survey on students’ attitudes toward confidence, success in college, social networking, student engagement, access and affordability, etc.
- Case studies or conceptual articles providing institutional perspectives on process development and maintenance methodology at colleges or universities.
- Case studies or conceptual articles addressing issues such as the role of faculty and administrators in quality systems.
- Case studies, research studies, or conceptual articles focusing on accreditation issues.
- Case studies demonstrating best practices using the Baldrige Education Criteria for Performance Excellence, including experience and recommendations for successful implementation.
- Case studies, research studies, or conceptual articles on scholarship of teaching, enhancing student learning, learning outcomes, learning outcomes assessment, student retention, best practices for using technology in the college classroom, etc.

In particular, we are looking for articles on the following topics: using assessments for continuous improvement and accreditation, showing how use of the Baldrige framework can increase student success, increasing engagement and quality of learning through lecture capture and other technologies, dealing with rising costs without jeopardizing learning, new programs for helping graduates gain employment, and merging research with practice (action inquiry).

Manuscript Review Process

We log all article submissions into a database and delete all references to you. These “blinded” versions then go to the editorial review team for comments and recommendations. The review process takes approximately two months during which time the reviewers advise the editor regarding the manuscript’s
suitability for the audience and/or make suggestions for improving the manuscript. Reviewers consider the following attributes:

1. **Contribution to knowledge**: Does the article present innovative or original ideas, concepts, or results that make a significant contribution to knowledge in the field of quality in higher education?

2. **Significance to practitioners**: Do the reported results have practical significance? Are they presented clearly in a fashion that will be understood and meaningful to the readers?

3. **Conceptual rigor**: Is the conceptual basis of the article (literature review, logical reasoning, hypothesis development, etc.) adequate?

4. **Methodological rigor**: Is the research methodology (research design, analytical or statistical methods, survey methodology, etc.) appropriate and applied correctly?

5. **Conclusions and recommendations**: Are the conclusions and recommendations for further research insightful, logical, and consistent with the research results?

6. **Readability and clarity**: Is the article well organized and presented in a clear and readable fashion?

7. **Figures and tables**: Are the figures and/or tables used appropriately to enhance the ability of the article to summarize information and to communicate methods, results, and conclusions?

8. **Organization and style**: Is the content of the article logically organized? Are technical materials (survey scales, extensive calculations, etc.) placed appropriately? Is the title representative of the article’s content?

9. **Attributions**: Are the sources cited properly? Are attributions indicated properly in the reference list?

You should use these attributes as a checklist when reviewing your manuscript prior to submission; this will improve its likelihood of acceptance.

There are three possible outcomes of the review process:

- **Accept with standard editorial revisions**: In this case, the content of the article is accepted without requiring any changes by you. As always, however, we reserve the right to edit the article for style.

- **Accept with author revisions**: An article in this category is suitable for publication but first requires changes by you, such as editing it to fit our length requirements. We provide specific feedback from our reviewers to guide the revision process. We also assign a tentative publication date, assuming you will submit the revised article by the deadline.

- **Decline to publish**: Occasionally articles are submitted that do not fit our editorial scope. In these situations, we may provide you with suggestions for modifying the article to make it more appropriate to our publication, but we do not assign a tentative publication date.

Please note that after articles are edited for publication, we return them to you to approve the technical content. A response may be required within 48 hours or the article may be held over for a subsequent issue.

Articles that appear to be advertising or don’t fit the general topics addressed by *Quality Approaches in Higher Education* will be rejected without receiving peer reviews.

**Helpful Hints**

1. **Articles should emphasize application and implications.**
   - Use the early paragraphs to summarize the significance of the research.
   - Make the opening interesting; use the opening and/or background to answer the “so what?” question.
   - Spell out the practical implications for those involved in higher education.
2. Detailed technical description of the research methods is important, but not necessarily of interest to everyone.

3. Throughout the article, keep sentence structure and word choice clear and direct. For example, references should not distract from readability. As much as possible, limit references to one or two per key idea, using only the most recent or most widely accepted reference.

4. Avoid acronyms and jargon that are industry- or organization-specific. Try not to use variable names and other abbreviations that are specific to the research. Restrict the use of acronyms to those that most readers recognize. When acronyms are used, spell them out the first time they are used and indicate the acronym in parentheses.

5. Our reviewers and readers usually view articles that include reference to your proprietary products or methods as advertising. Although we encourage you to share personally developed theories and application approaches, we ask that you refrain from using our publication as a marketing tool. Please take great care when including information of this nature in your article.

6. If the article cites cost savings, cost avoidance, or cost-benefit ratios, or provides the results of statistical evaluations, include an explanation of the method of calculation, along with any underlying assumptions and/or analysis considerations.

7. When submitting an article that includes survey data, include the complete survey instrument. We may make the entire survey available online.

8. Our staff does not have the means to compile references or verify usage permissions; therefore, it is important for you to provide all that information with your article, including written letters of authorization when appropriate. Plagiarism is a rapidly growing crime—particularly due to the use of information from the Internet. Please help yourself, and us, to maintain professional integrity by investing the time necessary to verify your sources and to obtain and document all necessary permissions. Information on our requirements for documenting references, along with specific examples, is included at the end of these guidelines.

**Submission Format**

1. We accept only electronic submissions in Microsoft® Word® format. Send electronic copies of articles to qahe@asqedu.org. Also please include an abstract of 150 words or less for each article. Include all of your contact information in a cover letter or your email message.

2. Tables should be included at the end of the article and must be in Microsoft Word. Each table must be referenced in the article and labeled, such as “Table 1: Graduation Rate by Major.” Do not embed .jpg, .tif, .gif, or tables in other similar formats in your article.

3. Drawings and other illustrations should be sent in separate Microsoft® PowerPoint® or Microsoft Word files; each item should be included in a separate file. All drawings and other illustrations must be referenced in the article, must be labeled, such as “Figure 1: Pareto Analysis of Student Participation in Department Activities.” Please do not use other software to generate your drawings or illustrations. Also, please do not embed .jpg, .tif, .gif, or drawings or illustrations in other similar formats in your article.

4. We can use photos if they enhance the article’s content. If you choose to submit a photo with your article, it must be a high-resolution .jpg or .tif (at least 300 dpi and at least 4” by 6” in size). We
cannot enlarge photos and maintain the required resolution. Photos should be sent in separate files and referenced in the article. Photos should be accompanied by a complete caption, including a left-to-right listing of people appearing in the photo, when applicable. Do not include any text with the photo file.

5. Also submit a separate high-resolution electronic photo (at least 300 dpi) for each author. Author photos should be at least 1" by 2". Author photos should have a plain background, and the author should be facing toward the camera.

6. Please include a 75- to 100-word biography for each author, mentioning the place of employment, as well as including a telephone number, Web site, and/or email address. If you have published books within the past five years, we encourage you to include the names of one or two books. We do not have space to mention articles, speech titles, etc.

Copyright Transfer

Prior to publication, you must sign a form affirming your work is original and is not an infringement of an existing copyright. Additionally, we ask you to transfer copyright to ASQ. The copyright transfer allows you to reproduce your article in specific ways, provided you request permission from ASQ and credit the copyright to ASQ. The transfer also allows ASQ to reproduce the work in other publications, on its Web site, etc.

If you use materials from other works in your articles (other than standard references), you must obtain written permission from the copyright owner (usually the publisher) to reprint each item of borrowed material. This includes any illustrations, tables, or substantial extracts (direct quotations) outside the realm of fair use. Submit these permission letters with the article. Articles cannot be published until copies of all permission letters are received.

For example, an article includes a PDSA illustration from a book. The permission statement would include: Figure 1 is from Nancy R. Tague’s *The Quality Toolbox, 2nd ed.*, ASQ Quality Press, 2005, page 391. This permission statement would appear in the caption just below the PDSA figure.

References

One of the most common errors we’ve observed with submitted articles is improper referencing. Two problems occur most frequently: information included without proper attribution in the references and formatting that does not meet our style requirements. The information in this section is intended to ensure your references adhere to our standards.

*Quality Approaches in Higher Education* uses its own reference style. All references should be consecutively numbered in the body of the text, using superscripts, and a matching number, also using superscripts, should appear in the references section at the end of the article. Do not include periods with the numbers or spaces preceding or following the numbers. If multiple references are associated with a particular citation, list each separately (do not show a range). For example, “The engineering department modified its program and created an integrated freshman curriculum2,3 to promote a comprehensive learning environment that includes significant attention to student communication skills.” Use a comma to separate the numbers, but do not include a space after the comma. Please do not use Microsoft Word endnotes or footnotes; also, please do not include citations in the body of the text, such as is used for APA style.
Examples

TYPE: Book, one author:

TYPE: Book, two authors:

TYPE: Magazine/journal article, one author:

TYPE: Magazine/journal article, two authors:

TYPE: Magazine/journal article, no month or year given, only volume and number:

TYPE: Web site articles:

TYPE: Conference proceedings:

Tips

• We use commas to separate segments of the reference information, not periods.
• Authors’ names always appear with the first name followed by the last name.
• The names of books, magazines, newsletters, and journals are italicized.
Author Guidelines: *Quality Approaches in Higher Education*

- Double quotation marks are used around the names of magazine, newsletter, and journal articles and conference proceedings’ titles. Punctuation marks fall inside the quotation marks in almost every case.
- It’s not necessary to include the city with the publisher’s name.
- When inserting the reference numbers in the body of the text, use the “superscript” function in Microsoft Word. Do not include a period behind the reference number or a space before or after the reference number, as shown below:
  
  *Correct:* Text in body of the article\(^1\)
  *Incorrect:* Text in body of the article\(^1\) 
  *Incorrect:* Text in body of the article \(^1\)

- When inserting the reference number in front of the reference information in the list at the end of the article, use the standard font size and format. Do include a period behind the reference number and a space after the period, as shown below:
  
  *Correct:* 1. Reference information
  *Incorrect:* 1Reference information
  *Incorrect:* 1 Reference information
  *Incorrect:* 1. Reference information

**Summary**

Thank you for considering having your article published in *Quality Approaches in Higher Education*. We look forward to reviewing your manuscript. Please feel free to contact Dr. Cindy Veenstra at qahe@asqedu.org if you have any additional questions.
Can you think critically about what you read?

Are you able to express yourself clearly and concisely.

Do you have expertise in quality approaches for higher education?

Are you willing to volunteer your time to help improve your profession?

If you can answer “Yes” to each of these questions, then Quality Approaches in Higher Education invites you to become a member of its Review Board. As a reviewer, you will be expected to maintain a standard of high quality for articles published in this journal and help build its reputation for excellence.

To become a reviewer, please complete the application on the next page and send it with a copy of your curriculum vitae to Dr. Cindy Veenstra at qahe@asqedu.org. Your application will then be reviewed by the editorial team and you will be notified in approximately 60 days if you have been accepted as a reviewer. Following acceptance to the Review Board, you will become part of the pool of reviewers available to evaluate articles submitted for publication. The frequency of your review assignments will depend on the number of articles submitted and the number of reviewers with the expertise needed to critically evaluate each article.

Once assigned to review an article, you will be emailed that article, which will have been “blinded” to remove information about the author(s) to assure your impartiality. Along with the article you will be sent detailed review instructions and the review form, itself. As you critically read the assigned article, your primary focus will be on the article’s content, not its style-related issues such as grammar, punctuation, and formatting. The editorial team is charged with assuring that all style-related issues are resolved in accordance with ASQ’s modified-AP style guide prior to publication. Your task is to provide ratings and detailed comments in nine content-related categories, plus an overall rating which reflects your recommendation for article disposition. You will be given approximately three weeks to return your completed review form for each article.

Article disposition will be determined by the editorial team based on input from the reviewers. In cases where a revision is recommended, detailed instructions will be provided to the author(s) using reviewer comments. Revised articles will be evaluated by the editorial team for compliance to the improvement recommendations, and one or more of the original reviewers may be asked for input as part of this process.

We look forward to receiving your application to become a reviewer.
Call for Reviewers: *Quality Approaches in Higher Education*

**NAME**

**ADDRESS**

**E-MAIL ADDRESS**

**PHONE NUMBER**

**EMPLOYER**

**JOB TITLE**

**DEPARTMENT**

**AREAS OF EXPERTISE**

Select all that apply

- Scholarship of Teaching
- Best Practices
- Enhancing Student Learning
- Technology in the Classroom
- Accreditation

**Research Methodology**

- Survey Design, Analysis, and Interpretation
- Research Statistics
- Organizational Research

**Other (please specify)**

- ________________
- ________________

**PERSPECTIVE**

Select all that apply

- Faculty
- Administration
- Student Services
- Consultant to Higher Education

- Student
- Undergraduate Education
- Graduate Education
- International
- Multi-site Institution

Applicant’s Curriculum Vitae must be included with this application.

Email to Dr. Cindy Veentra at qahe@asqedu.org.