

National Institute for Learning Outcomes Assessment

December 2009

THREE PROMISING ALTERNATIVES FOR ASSESSING COLLEGE STUDENTS' KNOWLEDGE AND SKILLS

Trudy W. Banta, Merilee Griffin, Teresa L. Flateby, and Susan Kahn

Foreword by Jillian Kinzie



Occasional Paper #2

learningoutcomesassessment.org

ABOUT THE AUTHORS

Trudy W. Banta

Trudy W. Banta is Professor of Higher Education and Senior Advisor to the Chancellor for Academic Planning and Evaluation at Indiana University–Purdue University Indianapolis. Recipient of 8 national awards for her work, she has developed and coordinated 22 national conferences and 15 international conferences on the topic of assessing quality in higher education. She has given addresses at conferences and/or consulted with faculty and administrators on college and university campuses in 47 states and 26 other countries. Dr. Banta has edited 16 published volumes on assessment, contributed 27 chapters to published works, and written more than 200 articles and reports. She is the founding editor of *Assessment Update*, a bimonthly periodical published since 1989 by Jossey-Bass.

Merilee Griffin

Merilee Griffin, Ph.D. is President of Collaborative Online Assessments (www.coassess.org), a nonprofit organization dedicated to research and advancement of COA's. She currently serves as Research Coordinator for the Tier I Writing Assessment Project at the Center for Writing in a Digital Environment at Michigan State University and consults on other projects.

Teresa L. Flateby

Teresa Flateby, Ph.D. with over fifteen years of assessment experience, recently retired as the Director of Assessment at the University of South Florida (USF) to establish her own consulting firm. She has facilitated numerous assessment workshops at the national level and has provided writing and thinking assessment consultation for a variety of higher education institutions.

Susan Kahn

Susan Kahn, Ph.D. is Director of Institutional Effectiveness at Indiana University–Purdue University Indianapolis (IUPUI), leads the university's electronic portfolio initiative, and holds adjunct faculty appointments in IUPUI's University College and School of Liberal Arts. She writes and presents widely on the use of ePortfolios for faculty, students, and institutions.

CONTENTS

Abstract	3
Foreword	4
Three Promising Practices	5
Introduction	5
Involving Faculty in Authentic Assessment	6
Electronic Portfolio Assessment	7
The CLAQWA Rubric System.....	12
Online Assessment Communities	18
Conclusion	24
References	25
NILOA	27
National Advisory Panel	27
NILOA Mission	27
Occasional Paper Description	27
About NILOA	28
NILOA Staff	28
NILOA Sponsors	28

The ideas and information contained in this publication are those of the authors and do not necessarily reflect the views of Carnegie Corporation of New York, Lumina Foundation for Education, or The Teagle Foundation.

reflect curiosity challenge create achievement connection se
innovation success ingenuity intellect curiosity challenge knowle
understand communicate curiosity challenge create achievem
reflection knowledge accountability connection self-reflection edi
ingenuity intellect curiosity challenge educate innovation success
connection self-reflection educate action understand communic
communicate listen learn access quality action educate action und



ate action understand communicate listen learn access quali
knowledge accountability connection self-reflection educate actio
reflection curiosity challenge create achievement connection sel
rstand communicate listen learn access quality innovation succes
t curiosity challenge create achievement knowledge accountabili
llenge create achievement connection self-reflection understand
cate listen learn action understand communicate listen learn acces

ABSTRACT

Three Promising Alternatives for Assessing College Students' Knowledge and Skills

Educators and policy makers in postsecondary education are interested in assessment processes that improve student learning, and at the same time provide institutional data that may be used to demonstrate accountability. This paper discusses three promising alternatives that afford the kinds of authentic, information-rich, meaningful assessments that are essential for improving student learning, and at the same time provide data for public reporting. First, ePortfolios offer an in-depth, long-term view of student achievement on a range of skills and abilities as opposed to a quick snapshot based on a single sample of learning outcomes. Second, a system of rubrics used to evaluate student writing and depth of learning has been combined with faculty learning and team assessments, and is now being used at multiple institutions. Third, online assessment communities link local faculty members in collaborative work to develop shared norms and teaching capacity, and then link local communities with each other in a growing system of assessment. These authentic and valid assessment approaches must be developed and promoted as viable alternatives to scores on single-sitting, snapshot measures of learning that do not capture the difficult and demanding intellectual skills that are the true aim of a college education.

knowledge accountability connection self-reflection educate action understand communicate listen learn access quality innovation success ingenuity intellect curiosity challenge create achievement connection self-reflection educate action understand communicate listen learn access quality innovation success ingenuity self-reflection educate action understand communicate curiosity challenge create achievement self-reflection curiosity challenge create achievement connection self-reflection communicate listen learn access quality innovation success ingenuity intellect curiosity challenge educate innovation success connection self-reflection educate action understand communicate listen learn access quality action educate action understand communicate listen learn action understand communicate listen learn access quality action educate action understand communicate listen learn action understand communicate listen learn access

FOREWORD

Every college or university must decide how to most effectively assess student learning outcomes for institutional improvement and accountability. As reported in NILOA's first report, *More Than You Think, Less Than We Need: Learning Outcomes Assessment in American Higher Education* (2009), most institutions use a combination of assessment tools and approaches that vary depending on what the data are intended to represent and how the results are to be used. Two thirds of institutions are using three or more approaches, including nationally standardized tests or surveys, local surveys, and authentic assessments such as portfolios and rubrics. In addition to measures of specialized knowledge, the most common approaches to assessing learning outcomes at the program level are student portfolios, other performance assessments, and rubrics. The latter three assessment approaches are often described as authentic assessment measures.

In this paper, assessment experts Trudy Banta, Merilee Griffin, Teresa Flateby, and Susan Kahn describe the development of several promising authentic assessment approaches. Such performance assessments typically ask students to generate rather than choose a response to demonstrate what they have learned, providing a holistic picture of learning gains over time. These approaches are valued by faculty because they tend to be flexible, can be closely aligned with teaching and learning processes, and represent some of students' more meaningful educational experiences. However, their use is less widespread than simple survey methods, in part due to questions about viability and -- in particular -- the cost of widespread use of student portfolios, the feasibility of validating locally-developed rubrics, and the challenges inherent in aggregating results to allow for meaningful cross-institutional comparison.

To increase the application and use of authentic assessments in higher education, the following concerns must be addressed:

1. How can institutions reasonably collect and organize authentic measurements of students' knowledge and skills?
2. To what extent can results from individual measures of student learning, such as portfolios and rubrics, be used to inform institutional improvement activities and to make valid judgments about institutional performance?
3. Can faculty within and across institutions develop shared standards for using authentic measures to evaluate student learning for purposes of accountability and improvement?

Banta, Griffin, Flateby, and Kahn address these concerns by drawing on their rich assessment experience to illustrate with institutional examples how authentic assessment approaches can be used to document student learning and guide institutional improvement efforts. In addition to discussing the strengths and limitations of each approach, the paper describes instances in which portfolios, common analytic rubrics, and online assessment communities are being used effectively to create common standards and expectations for student learning and for demonstrating accountability.

The authors present a persuasive case for authentic assessment as both an approach to yield deeper understandings of individual student learning and an efficient and effective method to assess learning outcomes and demonstrate institutional accountability. By highlighting concrete examples of good practice, they illustrate that authentic assessments hold substantial potential for integrating assessment practice and pedagogy that can foster higher levels of student learning and involve more faculty, staff, and students in meaningful appraisals of learning outcomes.

Jillian Kinzie
Associate Scientist
Indiana University Center for Postsecondary Research
NILOA Project Staff



Three Promising Alternatives for Assessing College Students' Knowledge and Skills

Trudy W. Banta, Merilee Griffin,
Teresa L. Flateby, and Susan Kahn

Introduction

Around the globe, educators and policy makers alike are increasingly interested in finding methods to assess postsecondary students' knowledge and skills, in part so that they can compare institutions' effectiveness in advancing student learning. In this paper we describe three assessment systems that promise authentic, comprehensive, and useful measurement of students' knowledge and skills for purposes of providing direction for institutional improvement as well as for demonstrating accountability.

Developing assessment systems must begin with the identification of some common expectations for what students should know and be able to do when they complete a program of study. Tuning, a component of the Bologna Process, has brought together faculty, students, and employers from across Europe in fields as diverse as business, chemistry, education, and history to identify both subject-specific competences expected of graduates and important generic skills such as communication and leadership (<http://tuning.unideusto.org/tuningeu/>). The Lumina Foundation has funded a pilot project in three states—Indiana, Minnesota, and Utah—to see if tuning might work in the U.S. The Quality Assurance Agency for Higher Education in the U.K. is beginning to investigate ways to develop explicit and comparable statements about achievement standards for tertiary graduates. Members of the Australian Universities Quality Agency are a bit farther along, having drafted a set of guidelines for setting and monitoring such standards (AUQA, 2009).

While progress is being made in identifying common learning outcomes across institutions and even across countries, development of measures to assess student achievement of these outcomes is in its infancy. This is not surprising since constructing assessment methods that are both reliable and valid is a long, difficult, and expensive process. The Assessing Higher Education Learning Outcomes project of the Organisation for Economic Cooperation and Development (OECD) seeks to pilot test measures of certain generic skills as well as measures of subject-specific knowledge and skills in engineering and economics in a few institutions in each of 10 countries (www.oecd.org/). In the U.S. the Voluntary System of Accountability contains a list of three commercial instruments designed to provide a snapshot in time of college students' skills in written communication, analytic reasoning, and critical thinking (www.voluntarysystem.org/). And the Association of American Colleges and Universities (AAC&U) is developing a set of VALUE (Valid Assessment of Learning in Undergraduate Education) rubrics that may be used to assess a wide range of student performances—written, visual, and oral—in each of 15 areas, including written communication, oral communication, creative thinking, teamwork, and ethical reasoning (see <http://www.aacu.org/value/>).

Developing assessment systems must begin with the identification of some common expectations for what students should know and be able to do when they complete a program of study.

With the exception of the VALUE project, relatively few faculty in the U.S. have been involved in the national conversations about the need to agree on common learning outcomes across institutions and to develop measures that facilitate comparisons of institutional effectiveness. This seems almost incomprehensible since it is faculty who determine students' learning experiences and who continuously assess students' learning. In addition, faculty alone can make improvements in curriculum, instruction, and student support services if assessment data suggest the need for such changes.

Involving Faculty in Authentic Assessment

We **must** involve faculty in the difficult work of developing methods of assessment that demonstrate institutional accountability. And when we seek their counsel, we learn that they prefer authentic measures—asking students “to do real-life tasks, such as analyzing case studies with bona-fide data, conducting realistic laboratory experiments, or completing internships” (Suskie, 2009, p. 26). Faculty want more than a snapshot in time of student learning, which might be taken on a very bad day in the life of a particular student. Faculty who have been consulted about assessment for accountability are interested in using the papers, projects, exams, and simulated or actual professional practice opportunities they assign in the process of stimulating learning—both to assign grades to individual students and to demonstrate accountability by aggregating assessments across students in a class, courses in a major, and courses at an institution. The VALUE rubrics illustrate one method for facilitating such aggregation.

Outcomes, pedagogy, and measurement methods must all correspond, both for summative assessment (demonstrating that our students have achieved certain levels) and formative assessment (improving student learning, teaching, and programs). To determine if students can think critically by evaluating the credibility of claims and the logical strength of arguments, evidence revealed in written, oral, or visual communication is required. Furthermore, assessment evidence must reflect the level of complexity and detail of real-life tasks, including locating information; evaluating the credibility of sources; understanding multiple viewpoints; synthesizing information from various sources; creating complex explanations, solutions, or theses that capture multifaceted realities; and integrating source material into an original work.

Faculty in hundreds of departments and programs across the country are working on such measures and have developed assessments that have high contextual validity and produce immediate, direct results in improving teaching and curricula. They are motivated to spend time creating rubrics because they know that rubrics will save them time in grading papers and in helping students understand the grading process. They also want to know if new approaches work better than previous ones and appreciate the role of assessment data in making these judgments. What has been missing to date is an answer to the demands for cross-institutional comparisons that require the merging of assessment data from local, faculty-led initiatives into a system of larger units where the assessment of learning outcomes has meaningful, operationalized, and transparent standards.

In this paper we describe three promising methods for providing authentic evidence of student learning. These approaches are attractive in part because they are components of the classroom context. They are embedded in course assignments, count toward students' final grades, and thus encourage students

We must involve faculty in the difficult work of developing methods of assessment that demonstrate institutional accountability. And when we seek their counsel, we learn that they prefer authentic measures . . .

to work conscientiously. When students take responsibility for doing their best work and do so over time and at their convenience, they are able to demonstrate more of what they know and can do. These authentic assessment methods furnish a more complete picture of what students have learned in college than many other assessment options. In addition, they give students the opportunity to demonstrate their own unique strengths and interests. Importantly, we also suggest a means for institutions to combine assessment methods, building a shared set of standards and expectations for demonstrating accountability. In fact, we describe instances in which this is already happening.

Each of the three approaches is described in the following format:

- Background
- Description of the method
- How the method has been used
- Strengths and limitations of the method

First, Susan Kahn introduces electronic portfolios, an increasingly popular form of authentic assessment that enables students to collect and reflect on evidence of what they are learning based on their own assignments and actual experiences in college. Next, Teresa Flateby tells us how faculty at the University of South Florida (USF) and other institutions are using rubrics to score student writing and critical thinking in various disciplines, an approach that might be applied to the artifacts in electronic portfolios. Finally, Merilee Griffin offers details about establishing assessment communities that could begin with local assessments like USF's and then be expanded to multiple groups of peer institutions.

Electronic Portfolio Assessment

Simply put, an electronic portfolio (ePortfolio) is “a digitized collection of artifacts, including demonstrations, resources, and accomplishments, that represent an individual, group, or institution” (Reese & Levy, 2009, p. 2). Student ePortfolios in higher education typically include reflection, along with the “digitized collection” of representative work, and can serve a range of pedagogical and assessment purposes. They can also provide students with the means to present their work in an accessible electronic format to potential employers and graduate schools. The promise of ePortfolios to support student development and employment as well as authentic assessment probably explains their increasingly widespread adoption by colleges and universities across the U.S. and around the world.

While the concept of authentic or performance-based assessment has been with us for some time, ePortfolios are still in their infancy, with very few examples of longstanding campus use. They represent a convergence of expanding technological capacities with emerging ideas and findings about how students develop and learn, what pedagogies most effectively support student intellectual growth, and how assessment can contribute to both improvement and accountability. These ideas and findings include:

- *A growing understanding of the importance of student engagement to successful learning* (Kuh, 2008; Kuh, Kinzie, Schuh, Whitt, & Associates, 2005). Well-designed ePortfolio programs require students to take an active role in selecting work and developing presentations that reflect their intellectual growth and mastery of key learning outcomes.

These authentic assessment methods furnish a more complete picture of what students have learned in college than many other assessment options. In addition, they give students the opportunity to demonstrate their own unique strengths and interests.

Through guided reflection, students gain understanding of their own learning processes and improve their ability to critique themselves. Portfolios can encourage students to take greater ownership of their learning and of their intellectual and professional development (Yancey, 2009).

- *A shift in focus from the teacher to the learner* (Barr & Tagg, 1995; Tagg, 2003) and a related emphasis on designing coherent, integrated curricula and learning experiences (Eynon, 2009; Huber, Hutchings, Gale, Miller, & Breen, 2007; Schneider, 2008). Selection of ePortfolio artifacts and reflection on them enable students to participate actively in integrating their learning across disciplines, semesters, and in- and out-of-class experiences (Cambridge, 2009; Eynon, 2009; Hamilton & Kahn, 2009).
- *An emphasis on outcomes—what students know and are able to do as opposed to the traditional emphasis on “seat time” and credits accumulated* (Barr & Tagg, 1995; Tagg, 2003). Portfolios contain authentic evidence of learning that students produce in the everyday course of their studies, both in and out of the classroom (e.g., in internships, independent research projects, and other “real-world” experiences). The capacity of digital environments to accommodate work in multiple media, including graphical and video formats, enables students to include actual performances that demonstrate the application of their learning (Lorenzo & Ittelson, 2005). For example, an informatics student can include multimedia projects she has designed. A teacher education student can include video clips from classes he has taught.
- *Interest in authentic, performance-based forms of evaluation and accountability.* Portfolio assessment is based on representative samples of student performances and on student reflection and self-analysis, using rubrics appropriate to the discipline and/or desired learning outcomes (Banta, 2007). Some ePortfolio software incorporates assessment capabilities that enable outcomes data to be aggregated and sorted for purposes of both improvement and accountability (Lorenzo & Ittelson, 2005).

Supporting these new approaches and emphases is the now almost universal use of e-learning environments, such as course and learning management systems, in higher education. Designers of these environments, whether commercial, homegrown, or open source, are increasingly seeking to incorporate capacities for rich learning experiences. Many of these environments now include some form of ePortfolio. Approximately 40 commercial vendors are currently marketing ePortfolio software in the U.S. (T. Batson, personal communication, April 3, 2009).

Description of the Method

The principal argument for portfolio assessment rests on validity: By incorporating samples of the actual work students produce as they move through the curriculum, ePortfolios demonstrate the learning outcomes that faculty intend students to master (Banta, 1999, 2007). Effective use of ePortfolios for formative and summative assessment thus requires that they be woven into the curriculum and co-curriculum to capture ongoing student work. Ideally, students contribute work and reflections to their ePortfolios over the entire course of their undergraduate studies or at least the study of their major.

The principal argument for portfolio assessment rests on validity: By incorporating samples of the actual work students produce as they move through the curriculum, ePortfolios demonstrate the learning outcomes that faculty intend students to master (Banta, 1999, 2007).

At Indiana University-Purdue University Indianapolis (IUPUI), for example, some departments designate specific “portfolio courses,” even requiring that certain assignments be added to ePortfolios, to ensure ongoing use of the ePortfolio among faculty and students. The concept of portfolio courses is also used in some institution-wide ePortfolio programs, such as the one at LaGuardia Community College in New York City (Eynon, 2009). In addition to producing artifacts that can be incorporated into portfolios as evidence of learning, these portfolio courses may also include assignments that ask students to reflect on their learning across a number of courses. Other institutions and programs may simply provide students with an introduction to the portfolio in an early course and then ask them to submit a portfolio for evaluation at designated intervals. Programs also vary in terms of what portfolio materials are actually assessed. Assessed materials may include key common assignments, assignments selected by students, reflections, or some combination of these.

Programs also take different approaches to when, where, and by whom portfolio assessment is carried out. For example, in programs that include designated portfolio courses, assessment of portfolios, or sections of portfolios, may be conducted by the faculty members responsible for those courses as part of their ongoing responsibility for evaluating student course work. In other cases, portfolio assessment takes place outside the context of courses, at designated points in students’ progression through the curriculum. Some programs use external evaluators, in addition to program faculty, often in response to accreditation or licensure requirements.

The ePortfolio Initiative at Virginia Tech incorporates a set of questions for programs to consider as they plan for adoption of an ePortfolio. These basic questions need to be answered by any department or program undertaking an ePortfolio project. They include the following:

- When and from where will student work be collected?
- Are there specific assignments that correspond to specific goals [i.e., learning outcomes]? Is the same assignment provided in different classes by different faculty? Is there a grading rubric for this assignment that might help standardize grading?
- Who will evaluate student ePortfolios? Will there be external evaluators?
- Which faculty in your department/program must adopt an ePortfolio to ensure that the needed student data are collected?
- Is there a need to have departmental consensus before the goals of your ePortfolio project can be enacted? If so, how might you go about ensuring buy-in and adoption? (Watson, 2007, p. 7)

How the Method Has Been Used

The discussions of the Cognitive Level and Quality of Writing Assessment (CLAQWA) and of assessment communities that appear later in this paper provide a wealth of information and ideas about using rubrics to assess authentic student work. These approaches apply equally well to assessing the authentic work in student ePortfolios. In this section, therefore, we focus on models for operationalizing and institutionalizing ePortfolio assessment. For example, at IUPUI, the implementation strategy for the ePortfolio has evolved from an early, somewhat naïve, expectation that campus-wide adop-

In addition to producing artifacts that can be incorporated into portfolios as evidence of learning, portfolio courses may also include assignments that ask students to reflect on their learning across a number of courses.

tion would follow quickly once the technological tools were in place. The current approach is to work closely with academic departments and schools over several years to plan for implementation of the ePortfolio and then to provide intensive technological and faculty development support as academic units pilot-test the portfolio and move toward wider adoption.

Specifically, the IUPUI ePortfolio initiative funds two-year Integrative Department Grants (IDGs) that focus on the systematic integration of general education and discipline-specific learning outcomes throughout the department's curriculum. Project leaders and colleagues typically spend the first year of the grant on mapping program-specific and general education outcomes to specific courses in the curriculum and, in some cases, redesigning the curriculum to ensure that all outcomes are sufficiently addressed. This curriculum mapping exercise is crucial to answering the above question about when and from where student work will be collected. Planning also includes determining how ePortfolio tools will be structured within the campus's learning management system and creating guidance and reflection prompts for students within the portfolio environment. Finally, departments may also need to develop, refine, or select appropriate assessment rubrics. During the second grant year, departments pilot the ePortfolio and revise and refine plans, rubrics, guidance, and other aspects of their implementation model.

IUPUI faculty have found that successful ePortfolio implementation requires intensive and extensive faculty development and consultation. The IUPUI ePortfolio team works closely with staff in IUPUI's Center for Teaching and Learning to provide workshops on such topics as project planning, curriculum mapping, and rubric development and to assure that individual faculty members and departmental teams have ready access to one-to-one consulting resources. In short, the campus has learned key lessons about institutionalizing ePortfolio assessment on a large research campus: Provide incentives and support, move forward slowly and carefully, emphasize the needs of academic units, and develop examples of success among the most interested departments to spur demand among later adopters.

Similar lessons come from LaGuardia Community College's successful experience with ePortfolios. The director of that program, Bret Eynon, exhorts portfolio adopters to "be patient—this is a long-term process"; to think carefully about pedagogy, structure, and software; and to "build sturdy support structures" for students and faculty. The LaGuardia model includes "recursive assistance" for both groups, including faculty development seminars and dedicated labs staffed with trained student assistants to support student and faculty work on ePortfolios (Eynon, 2009, pp. 66-67). And from the University of Washington, which has also developed a large ePortfolio program, we hear comparable advice: Tom Lewis and Janice Fournier counsel campuses starting an ePortfolio initiative to "be relentlessly inclusive" in all aspects of developing and implementing ePortfolios; to "expect some failures"; and to offer "flexible, multifaceted support strategies for the technology and for doing thoughtful work with ePortfolios" (Lewis & Fournier, 2009, p. 132).

Strengths and Limitations of This Method

To ePortfolio enthusiasts, the advantages of ePortfolio assessment may appear to far outweigh the disadvantages. Nonetheless, for campus ePortfolio initiatives, it is important to anticipate and plan for the barriers and pitfalls that faculty may perceive or encounter as they begin to work with ePortfolio tools and concepts.

... it is important to anticipate and plan for the barriers and pitfalls that faculty may perceive or encounter as they begin to work with ePortfolio tools and concepts.

Advantages.

- The work assessed is drawn from students' ongoing class assignments and co-curricular work. As such, it represents the learning that faculty at the specific institution and in the specific discipline are actually teaching and helping students master.
- Student artifacts in ePortfolios represent work assigned in classes or produced in co-curricular activities in which students participate voluntarily. Research has established that most students are intrinsically motivated to succeed in their courses and co-curricular lives; therefore, portfolio assessment outcomes are likely to reflect the full extent of student learning and intellectual development in college.
- Portfolios capture complex higher-order learning outcomes because the artifacts they contain represent long-term, multistep, cognitively complex student performances and because portfolio assessments typically consider multiple pieces and types of work (Banta, 1999).
- Portfolios can support student engagement and personal and intellectual development. Research on ePortfolios indicates their potential to help students integrate and articulate their learning in ways that are personally and academically meaningful to them, as well as to external stakeholders (Cambridge, 2009; Day, 2009; Edwards & Burnham, 2009; Eynon, 2009; Hamilton & Kahn, 2009; Yancey, 2009).
- Portfolio assessments of key outcomes can be graded using rubrics, yielding numerical scores that are reasonably reliable (see first bullet point under "Disadvantages," below). Universities can provide the numbers that some stakeholders demand by aggregating these scores. Wide adoption of portfolio assessment across institutions could even produce comparative information (Banta, 2007).

Disadvantages.

- Increased contextual validity, an important strength of portfolios, often comes at the expense of reliability. However, well-designed and carefully tested rubrics, along with training, can minimize disagreement among portfolio assessors. Under these conditions, portfolios can be just as reliable as the essay components of nationally normed tests, which are scored in similar ways, i.e., using rubrics.
- Few campuses or departments have a sufficiently long history of ePortfolio assessment to generate meaningful longitudinal data or to provide models that others can adapt. Nevertheless, methods such as those in the following sections of this paper are being developed and offer promise.
- Electronic portfolios still utilize a young technology, and currently available software products rarely satisfy all user desires and needs. Most ePortfolio software is designed with one or two primary uses in mind and can be difficult to adapt to other purposes.

As noted above, ePortfolio assessment requires significant up-front planning and may necessitate substantive curriculum revision. This kind of work involves extensive faculty collaboration to achieve consensus on outcomes and criteria. Arguably, these "disadvantages" are also advantages of ePortfolio work, as is asserted in the section of this paper on assessment communities. But some faculty members may be unwilling to engage in these activities or may believe—accurately, in some cases—that these efforts will not be

ePortfolio assessment requires significant up-front planning and may necessitate substantive curriculum revision. This kind of work involves extensive faculty collaboration to achieve consensus on outcomes and criteria.

concretely recognized and rewarded in the traditional processes of promotion and tenure and determination of salary increases. Moreover, portfolio initiatives may founder because leaders fail to anticipate the difficulties of achieving wide faculty participation and consensus. The following sections on the CLAQWA rubric system and assessment communities suggest strategies for achieving faculty engagement.

The CLAQWA Rubric System

A second alternative to more traditional assessment approaches is the integration of assessment and pedagogy exemplified in the performance-based rubric system developed at the University of South Florida (USF), a system now employed in varying degrees at a number of institutions across the country. Faculty use this system—recently expanded into an online environment—to improve writing and critical thinking, to help students develop insight into their own strengths and weaknesses, and to assess programs.

Description of the Method

The Cognitive Level and Quality of Writing Assessment (CLAQWA) rubric was initiated in response to instructional assessment needs in a two-year, team-taught, general education learning community program at USF. Faculty used writing assignments to foster deeper learning about the course content and also to improve student writing. Consistent with research on writing assessment (Elliot, 2003; White, 1994), the grading criteria faculty used sometimes varied widely and at times were not expressed at all. To provide guidance for consistent grading, CLAQWA describes specific cognitive levels as well as the full range of writing skills across the disciplines. The cognitive portion of the scale, based on Bloom's (1956) Taxonomy of Educational Objectives—Cognitive Domain, and the writing portion of the scale, derived from commonly used writing handbooks, have evolved over the 12 years since CLAQWA's inception. Flateby and Metzger (1999, 2001), both involved in the learning community program, developed the writing scale iteratively with faculty and graduate student teams. Designed to be flexible for application across the curriculum, the 16-trait analytic rubric is jargon free and can provide formative as well as summative information.

After several studies comparing CLAQWA results with the state College Level Academic Skills Test (CLAST) scores, definitions for all five levels of student competence were supplied to enhance clarity and consistency and to achieve acceptable inter-rater reliability. After a number of iterations, the CLAQWA scoring team began scoring essays at agreement levels of 0.8 and above on each of the 16 traits (Table 1 includes a portion of the CLAQWA rubric). This change afforded the identification of specific weaknesses in student writing for formative assessment purposes.

The Cognitive Level and Quality of Writing Assessment (CLAQWA) rubric was initiated in response to instructional assessment needs in a two-year, team-taught, general education learning community program . . .

Table 1. Language: Contextual and Audience Appropriateness

Trait 12: Word Choice

Level

- 5 Vocabulary reflects a thorough grasp of the language appropriate to the audience. Word choice is precise, creating a vivid image. Metaphors and other such devices may be used to create nuanced meaning.
- 4 Vocabulary reflects a strong grasp of the language appropriate to the audience. Word choice is accurate.
- 3 Vocabulary reflects an inconsistent grasp of the language and may be inaccurate or inappropriate to the audience.
- 2 Vocabulary is typically inaccurate and inappropriate to the audience. Word choice may include vague, nondescriptive, and/or trite expressions.
- 1 Word choice is limited to vague, nondescriptive, and/or trite expressions and may include homonyms, errors, word choice inappropriate to the audience, and “thesaurus writing.”

Trait 13: Comprehensibility

Level

- 5 All sentences are clear and understandable.
- 4 The sentences are clear and understandable with rare ambiguities.
- 3 Most sentences are understandable but may include ambiguities.
- 2 Many sentences lack clarity and may misuse academic language.
- 1 Most sentences lack clarity and may misuse academic language.

Trait 14: Sentence Construction

Level

- 5 Clear and concise sentences vary, with the degree of complexity reflecting the audience and purpose.
- 4 Sentences vary, with the degree of complexity reflecting the audience and purpose.
- 3 Sentence variety is limited but attempts complex structure.
- 2 Complex structure is attempted without success and/or sentence structure is simplistic, but not throughout the text
- 1 Sentences are simple and repetitive.

Trait 15: Point of View

Level

- 5 Point of view is consistent and appropriate for the purpose and audience.
- 4 Point of view is appropriate for the purpose and audience, and a rare shift returns to the original point of view.
- 3 Point of view shifts occasionally, or may be consistent but inappropriate, for the purpose and/or audience.
- 2 Point of view is attempted, but shifts frequently.
- 1 Point of view is not established, confusing the reader.

(CLAQWA Online, Cross-Disciplinary. Retrieved from <http://claqwa.com>)

Shortly afterward, the rubric was modified for the peer review process. When the peer review rubric and accompanying guidelines were used in tandem, students' writing and thinking improved. Faculty members across the campus were offered the opportunity to have the process introduced in their classes. In departments as diverse as theatre and electrical engineering, faculty began using this approach with success and observed improvements in student writing (Flatey & Fehr, 2008).

Responding to needs identified through a faculty survey, the CLAQWA developers transformed the classroom and program assessment writing rubric into an online system (CLAQWA Online, <http://claqwa.com>) by displaying examples of student work for each level of each trait, including explanatory comments (see Figure 1 for a prototype of a screen shot of CLAQWA Online). The online system assists faculty, students, and assessment professionals to evaluate or provide feedback on student writing and thinking across the curriculum and to close the assessment loop. The online information helps students understand performance at each level and improve their writing on a specific trait.

Figure 1. Prototype of a CLAQWA Reasoning Screenshot

CLAQWA Online:
Building Your Thinking and Writing Skills

Example for Reasoning - Level 2

Assignment Description

Imagine that you have been given the opportunity to design and teach an "ideal course" in your major. Make sure to explain why the elements you discuss would be important to an academic audience. Your essay should address the following questions: What subject matter and method(s) of instruction will you include in your course? What type of learning will occur? (Examples: learning facts, analyzing concepts, applying skills or ideas to new situations, creating solutions to complex problems) What would you do as the teacher and what would you expect your students to do in return? What types of interaction will occur among members of the class? What do you expect the students to get out of the course and how would you know (assess) what was accomplished?

During high school, I was never a participant in a "perfect class." Therefore, I do not feel that I am fully qualified to comment on this type of classroom environment. As a freshman, I am also too inexperienced to offer criticism for college level classrooms. I do, though, hope to find a class that suits my educational goals. Thus, one class we all can look forward to participating in would be an effective English course.

In the best version of this ideal English that includes expectations, grade calculation, teachers or professors will leave these learning outcomes as a surprise. In this manner, students can easily follow the professors' lesson plan and incorporate course requirements into their schedules. Project objectives and attendance policies can be shared with students, as well as how technology will be integrated into the course. Therefore, students will be aware of what is expected of them.

All students in college need to gain confidence. Confidence will allow students to effectively analyze scholarly texts and compose insightful, written arguments. Being confident can not possibly make them better writers and readers. The professor would urge students to collaborate in small peer review sessions. By working together, they will clearly be able to work independently.

However, many of my reasons for creating an ideal English course are actually quite selfish. I have

Annotations:
 - This statement assumes all of the students in this class will look forward to participating in an effective English course.
 - Assumptions mark a flaw in logical thought.

In addition to enhancing the writing rubric, placing it online, creating a peer review process, and developing instructional examples, the developers added components for critical thinking. Two sources provided the foundation: the American Philosophical Association's (1990) *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction*, commonly called the "Delphi Report," and Bloom's (1956) Taxonomy of Educational Objectives—Cognitive Domain. A cross-disciplinary faculty team defined the traits of analysis, evaluation, and inference in each of the five CLAQWA levels. These new components formed a separate critical thinking rubric.

The CLAQWA rubric captures the contents of the VALUE Written Communication rubric, with the four levels of the VALUE rubric encompassing CLAQWA levels 2-5. Although the CLAQWA Critical Thinking rubric is more narrowly focused on Bloom's work and the American Philosophical Association's 1990 report, similarities exist between its theoretical underpinnings and those of the VALUE rubric.

How the Method Has Been Used

The CLAQWA system has been used online for peer review, grading, instructor feedback, and program assessment. Instructors use the rubric on paper or online to guide feedback and evaluation of their assignments, to communicate their expectations, and to engage students in the peer review process. The online format also provides for storage and document management. Using CLAQWA Online, students have a convenient way to conduct peer reviews and to view examples of each trait for planning or revising their writing.

Through the developers' conference presentations, articles, and online presence, faculty from other institutions have learned about CLAQWA and adopted or adapted the rubric for classroom or program assessment. Instructors at Coker College, Columbus State University, Eastern New Mexico State University, Mitchell Community College, Oakton Community College, and Voorhees College have used it. A recent application of the rubric at New College of Florida, a small liberal arts college, resulted in improved writing and student insight into writing strengths and weaknesses. The system can be used in secondary schools as well as colleges and universities. Calvary Christian High School in Clearwater, Florida has begun using it for the assessment and development of student writing.

The online system is most useful for program assessment. At USF a team of trained scorers evaluates student submissions that are faculty-developed assignments. Because results are aggregated, the assignments are reviewed to ensure similar levels of cognitive complexity and degrees of specificity—a necessary step if comparisons are planned. Due to the comprehensive nature of the CLAQWA rubric, knowledge of the assignment's content is necessary to produce valid and reliable scores. For these faculty-developed assignments, the instructor who created the assignment is invited to attend a portion of a scoring session to ensure accurate evaluation of the rubric elements pertaining to the content. When students' work is evaluated with CLAQWA Online, they receive scores on each of the rubric elements as well as embedded comments that identify problem areas (Figure 2 shows an example of this type of feedback). Using CLAQWA, student essays are assessed on each of the 16 skills or traits on a five-point scale. Instructors receive a brief report including aggregated feedback on their students' performance and a ranking of the strongest and weakest elements for use in formative assessment of their instruction (see Table 2). The report also shows the percentage of students reaching mean-

Using CLAQWA Online, students have a convenient way to conduct peer reviews and to view examples of each trait for planning or revising their writing.

ingful predetermined score points that represent “below desired,” “acceptable,” and “desired” performance levels. Similar reports can be provided for a department or a college. Although scores in the 3.5-4.0 range are desirable for students in exit classes, our students’ scores are typically lower. In addition, scores below 2.5 reflect remedial level writing, which should rarely occur in exit classes. It should be noted that scores do not indicated performance beyond this particular essay; in other contexts students may perform differently.

Figure 2. Sample of a CLAQWA Online Reviewed Paper

CLAQWA Online *Building your Thinking and Writing skills*
Cognitive Level and Quality Writing Assessment Instrument

U John Dewey, the “Public” and the Issues of Social Justice

University of South Florida (Tampa)
Philosophy recovers itself when it ceases to be a device for dealing with the problems of philosophers and becomes a method, cultivated by philosophers, for dealing with the problems of men.
John Dewey, “The Need for a Recovery in Philosophy”

U **Abstract**

The views of social justice **U** in vogue today, whether they hail from the **U** classical liberal tradition or the Hegel-cum-Marx, seem ill equipped in addressing current issues surrounding **U** disparate lines of thought are rooted in their **U** devotion to abstraction **U** some theoretical entity as the moral foundation of social justice. But, **U** postmodern critiques, these approaches have been hard-pressed to establish the universality (à la Rawls) of their foundations. One way to avoid this problem might be to take a meliorative approach toward social issues, one that does not rely upon philosophical abstractions. Fortunately, John Dewey cultivated just such an approach, later dubbed “instrumentalism,” in the late 19th and early 20th centuries.

In 1927, Dewey released *The Public and Its Problems*, a series of lectures that were the fullest expression of “instrumentalism” relevant to political thought. There, Dewey anticipated the need for moving from the nationalism of “great society” towards the cosmopolitanism of “the great community” and developed what he saw as the philosophical groundwork leading to the latter. On his view, the concept of a “public” is driven by problematic situations facing a particular group at a particular time, and thus “the public” remains an ill-defined and illusory notion. Dewey’s approach provides a means for the “public” to realize itself and move towards the establishment of a world community that is based upon “native human intelligence” rather than on philosophical abstraction. With this in mind, as I will argue, social justice moved out of the realm of the ideal and into the concrete, thus making it more achievable.

Comprehensibility-Level3-Most sentences are understandable but may include ambiguities.

Table 2 contains the skill/trait scores of students in junior level courses arranged in descending order by mean. Also included are percentages of these students reaching the 2.5 level and not exceeding 3.5. The data suggest that the strongest skills relate to presenting and maintaining a main idea, including sufficient quantity of details to develop the main idea, and fulfilling assignment requirements, although even for these skills, the percentage of students reaching the 3.5 level is lower than expected. Students’ writing (for this essay) is weakest in these traits: supporting the main idea in the closing and opening, reasoning, and quality of details used to support the main idea. In summary, more scores should be in the 3.5 and above range, and many fewer scores should be in the 2.5 and below range.

Table 2. Class Report of CLAQWA Scores, Spring 2008

CLAQWA Scores for Spring 2008			
Mean And Percentage Scoring 3.5 or Above and Below 2.5 (n=17)			
Element	Mean	Percentage 3.5 and above	Percentage below 2.5
Ideas are comprehensible	3.24	53%	0%
All assignment requirements are fulfilled	3.06	41%	18%
A main Idea is presented and maintained	3.06	12%	0%
Details are sufficient in quantity to develop main idea	3.03	12%	0%
Purpose is clear and specific	3.00	18%	6%
Appropriate audience(s) is (are) consistently addresses	2.97	18%	0%
Coherence devices are present and appropriate	2.97	24%	12%
Sentence construction varies appropriately	2.97	24%	0%
Paragraphs demonstrate unity	2.91	6%	6%
Word choice is accurate	2.91	18%	6%
Grammar and mechanics	2.88	24%	12%
Point of view is consistent	2.78	12%	18%
Details are sufficient in quality to develop main idea	2.76	0%	12%
Opening supports main Idea	2.74	6%	12%
Reasoning supports main idea	2.62	0%	18%
Closing supports main idea	2.56	0%	24%

Strengths and Limitations of This Method

The CLAQWA Online system has important advantages over some other approaches to assessment. First, it allows faculty to create their own assignments for assessment, which can be as challenging and complex as they desire. The system can be used with portfolios as well as single works, so it is compatible with either a course-level or curriculum-wide portfolio requirement. Also, because the work being assessed is required for course completion, students will be motivated to do well and their work will produce credible assessment results.

The CLAQWA system is rich with feedback when implemented well. Feedback provides a comprehensive view of student strengths and weaknesses, including both formative assessment to aid students in improving their work and summative feedback that may be used for reporting purposes. Since students receive detailed feedback and have access to the online examples, the system is integrated with pedagogy to support student learning, which makes it especially appealing to faculty. Additionally, the focus on learning through reflection that usually accompanies portfolios is supported by CLAQWA's emphasis on critical thinking.

Beyond these advantages for student learning, the CLAQWA system affords easy and reliable comparisons among institutions. The standards inherent in the rubric, the samples used to exemplify the standards, and the constant use of the rubric by many different scorers have established a shared under-

standing of what student work should look like at each level. In addition, the use of external scoring by CLAQWA teams, perhaps working in assessment communities as described in the next section, would not only serve to unify curricula but also would be cost effective.

The primary disadvantage of the system is that the degree of complexity entailed in the rubric must be supported by the assignments faculty create for students, and these assignments must be reviewed in advance to ascertain that they contain a sufficient level of cognitive complexity. In addition, training for the team of scorers demands time and attention.

Online Assessment Communities

The idea of assessment communities builds on the development of successful local assessments such as those just described at USF. Many departments in postsecondary institutions have already developed common curricula and scoring rubrics faculty use to focus and standardize their teaching and grading. The assessment community extends the value of these efforts in two ways: first, it implements these efforts as the foundation for continuing professional practice, and second, it provides for the gradual merging of local assessment communities with those of similar departments or institutions to form the basis of a national assessment process.

Background

The inspiration for assessment communities comes from Lave and Wenger's (1991) work on situated learning and Wenger's (1998) subsequent book on communities of practice. The great contribution of these two thinkers is their explanation of how learning occurs not as a process of acquiring facts and concepts but as a gradual induction into a community of practitioners, where novices work alongside "adepts"—observing their practice, learning the skills in a gradual progression from simple to complex, and internalizing the language and values of the profession. Furthermore, professional practice is not merely a matter of applying proven formulas to data; rather, most cases in law, medicine, education, and other professions present nontextbook dimensions, so the practitioner must apply not only specialized disciplinary knowledge but years of experience and expert judgment (Mintzberg, 2000).

When assessment is viewed as a professional practice, several shifts occur in the way we think about it. First, practitioners require learning and practice and will continue developing their expertise throughout a career. Second, assessment is central to the practice of teaching in the same way that diagnosis is central to the practice of medicine. And third, assessment is not a technical process that can be outsourced to minimally trained technicians; high levels of professional education and expertise, such as the CLAQWA team represent, are required to evaluate the complexities, multiple dimensions, and nuances of the work being assessed.

The literature of learning communities and communities of practice maps almost perfectly onto the practice of assessment. In their collegial discussions about the assessment of student work, faculty not only acquire knowledge from others but also clarify and extend their own thinking in the process of explicating it to others. Dialog brings assumptions and biases to the surface where they can be re-examined, modified, and nested within new understand-

When assessment is viewed as a professional practice, several shifts occur in the way we think about it.

ings, making way for new thinking. Members gain confidence by taking the risks of new learning with the support of others. Finally, the dialog produces a certain amount of “growth-producing stress” that motivates people to change (Bruffee, 1999).

Because higher level skills are complex and multifaceted, the procedures used to assess them must be equally complex and multifaceted. Simplistic measures fail to discern subtle but important discriminations in language and thought processes, features that are often hard to describe but are powerful in their total effect. In assessment communities, each person brings different perceptions and areas of expertise to the table. The varied backgrounds, education, specialties, and interests of faculty are nearly always a source of mutual learning. Like different pieces of a puzzle, these diverse contributions combine, creating a varied, deep, enriched, and finely tuned assessment.

Description of the Method

Assessment communities are comprised of faculty who engage in collegial dialog with each other about how student work should be assessed. Most teachers will be greatly relieved to hear that membership in an assessment community does not necessarily entail any meetings, as the whole enterprise can be conducted online. Each member is then free to work wherever there is high-speed Internet access, on whatever schedule is convenient. Work can be spread over a semester or a summer.

Through the process of scoring and discussing student work, teachers “norm themselves” to score consistently with each other, in the same way scorers of commercial writing tests are normed to produce reliable scores. The work being scored consists of real samples produced in course work, presenting the same range of student ability, variety of problems, complexity of tasks, and demands of particular contexts that exist in reality. The values and priorities of the student work closely match the values and priorities of the program, because the same group of people develops both the program and the assessments.

Group members could be faculty in a small department who teach a particular course or, for higher level cognitive skills like critical thinking, a cross-curricular group representing several departments. Members analyze and score samples of student work at an interactive Internet web site. The amount of student work required for participants to form common judgments will vary from one program to another, and some experimentation will be necessary to determine how many work samples are required for faculty to reach a high degree of accord. In a recent trial in a first-year composition program, 25 student papers worked well, although good results were obtained after 20 papers; more gains might have been realized if the project had continued (Griffin, 2009).

The goal of the community is to achieve an inter-rater reliability coefficient of 0.80, a level that establishes credibility for the assessment community’s work. In the trial involving teachers of composition mentioned above, raters achieved a correlation of 0.94 after 20 papers and 0.88 after 25 papers (Griffin, 2009). These levels of agreement were remarkably high given the fact that there was no control of the prompt or assignment, no rubric was used, the scoring scale was a challenging nine points instead of six, and faculty were not even asked to approximate each others’ scores.

Because higher level skills are complex and multifaceted, the procedures used to assess them must be equally complex and multifaceted.

Samples of student work should be selected to present the full range of ability typical of the student population, and they should also represent the variety of problems faculty often see in student performance. Depending on the length of each piece, the samples should be divided into small batches that are appropriate for faculty to view, analyze, score, and discuss in a week. In the composition trial discussed above, five papers were selected for each week, with participants spending five or six hours per week for five weeks reviewing them. The project could be condensed or extended depending on the wishes and needs of participating faculty. Timelines should be clarified at the outset of the project: How many samples are to be analyzed and in what time frames?

If a department or program already has a group of faculty with special expertise in assessment, these individuals may constitute a suitable “expert group.” The expert group develops the content of the web site in accordance with course or program curricula and learning outcomes. In concert with each other, members of the expert group decide on a score for each paper and on the features and qualities of each sample that support their judgment.

The assessment community’s analysis involves identifying the strengths of a student’s writing as well as its faults in extensive detail. Vague summary comments, such as “the student failed to relate her ideas to the thesis,” are not sufficient. The analysis must delve into a point-by-point, almost sentence-by-sentence, level of detail to be meaningful. All comments except those that introduce and summarize ideas should quote words, phrases, or other material from the paper to illustrate and support the point of the comment.

This last idea is very important. Since much of the problem in communicating about assessment stems from the broad definitions of terms such as “logical” and “relevant,” it is essential to *operationalize* such terms by referring to specific instances. There is a great deal of difference between a comment that consists only of the statement, “The writer failed to relate her ideas to the thesis,” and an explanation in which that statement is followed by this commentary:

Notice, for example, in paragraph 4 (p. 2), where she relates several statistics about deaths and injuries occurring in the workplace, but she concludes the paragraph without telling how those statistics relate to inadequacies of OSHA rules. She does the same thing in the following paragraph on p. 3, which contains a long quote from Congressional testimony, but never explains how that quote relates to the faults she finds with OSHA rules.

There’s another example on p. 6, where she narrates the story of Dan, the ironworker. It’s an effective rhetorical device, but there is no introductory or follow-up material that tells how better OSHA rules might have prevented his accident. She needs to explain in her own words how those pieces of evidence support her thesis.

The explicit examples cited in this analysis are what make assessment criteria understandable and meaningful to the degree that they can guide faculty in making precise judgments in assessing student work. In other words, the examples operationalize the criteria. If the student paper were drawn from a first-semester composition course, the examples would illustrate the range of work expected at that level. If the paper were drawn from a capstone course in public administration, we would expect the examples, language, and complexity of argument to be more sophisticated. Without specific exam-

Since much of the problem in communicating about assessment stems from the broad definitions of terms such as “logical” and “relevant,” it is essential to operationalize such terms by referring to specific instances.

ples illustrating the criteria, it is impossible for faculty to distinguish among different levels of student achievement with any degree of reliability.

The technology required by an assessment community is very simple: a welcome web page, some introductory material providing theoretical background, and simple directions for logging on and navigating. The working part of the site begins with an index page listing all the student works to be assessed and leading the participant through each step in the assessment process.

After logging on, the teacher downloads a sample of student work, then analyzes its strengths and weaknesses offline. When ready, the participant logs back on, enters a score on the “scoreboard,” and then joins an asynchronous discussion forum for that paper. All participants comment at length on the paper and respond to each others’ posts. The web site is configured so that a person must enter a score before viewing scores given by other people or joining the discussion. The continuing comparison of one’s own judgments with the judgments of others is the mechanism for bringing assumptions and biases to the surface and for broadening perspectives.

How the Method Has Been Used

Assessment communities by a different name were widespread in English and writing departments in the late 1970s, when many teachers were trained in the newly developed method of holistic scoring. Although quite favorably received, assessment communities were time intensive and expensive, and they disappeared with the budget cuts of the 1980s. With activity and asynchronous discussion forums online, however, assessment communities are convenient because participants can work anywhere and at any time. Assessment communities are also very low cost because the technology involved is so simple. The result is a highly valid measure of student learning because the work assessed is the same work produced in the classroom.

The first trial of this method was at a large community college where four faculty members prepared analysis and commentary on 25 student papers from first-year composition courses. The other participants were six teachers of 11th- and 12th-grade writing courses in the region. The purpose of this trial was to learn whether an online assessment community could produce deep learning about college writing that would enable high school teachers to prepare their students for postsecondary work. The results were quite promising. High rates of inter-rater reliability were reached, and raters agreed on six key dimensions of student writing. Most important, all 10 participants made favorable comments about their experience in the project and said they would recommend it to others or would do it again (Griffin, 2009).

Shared assessments in two community college programs illustrate the faculty development value of assessment communities. At Tompkins Cortland Community College, faculty teaching a social sciences capstone course have been collaboratively assessing student learning since 2003 in a process that has led to ongoing deliberations about goals and a yearly review of the common rubrics. Jeanne Cameron, a faculty member there, has written that the most important results have not been specific revisions to the rubrics but discussions about what is valued in student work. Sharing their various insights and value systems, many faculty members have developed keener and broader perceptions as they evaluate the capstone products. Such discussions have been especially important in this cross-disciplinary context, where different discipline-based styles and standards lead to differing opinions about student

Assessment communities are also very low cost because the technology involved is so simple. The result is a highly valid measure of student learning because the work assessed is the same work produced in the classroom.

work. Furthermore, variation of scores from one semester to another has led faculty to calibrate their assessments more finely and to work for greater consistency (Banta, Jones, & Black, 2009).

Faculty of the first-year writing courses at Lansing (Michigan) Community College (LCC) are now in their 13th year of departmental review of student writing portfolios for each of the two required semesters. In lieu of six hours of classroom teaching, faculty members are paid each semester to evaluate portfolios of students from other teachers' courses, and each portfolio is read by two scorers. Feedback to students becomes more meaningful when provided by outside scorers, and faculty efforts during both semesters are aimed at helping students develop their best possible portfolios for the departmental review. Each teacher fully understands how his/her students' work will be scored by colleagues according to departmental standards. Richard Reagan, chair of the LCC Department of Communications, reports that the process of shared portfolio assessment has generated discussion about what is valued in student writing, and participation in such discussions has increased faculty knowledge about writing and grading (R. Reagan, personal communication, August 10, 2009).

Although neither of these assessment communities functions entirely online, both use Internet communications to some extent. At Tompkins Cortland, notes from an annual face-to-face meeting are distributed via e-mail, and the nine faculty members then respond online to each other's comments, with the ensuing conversation often resulting in revisions to the rubric. At LCC, much of the process occurs online, and developers are investigating online systems for the electronic exchange of portfolios.

Up to this point the focus of our description of assessment communities on work at the local level has been appropriate because local communities establish high validity for assessments. Local communities can create and maintain standards that teachers use in their own classrooms to provide better feedback to students, to guide their teaching, and to grade precisely. Standing alone, however, assessment communities do not address the need for reporting learning outcomes in a way that can be easily understood by the public and used for institutional accountability.

A second phase of assessment communities would begin after local online communities are well established. Each local community would seek others with similar curricula, missions, and students. At first they would view each other's assessment web sites, then begin trial joint assessments in which members exchange perceptions and values about student work, identify commonalities, and negotiate differences. Finally, the groups would merge their assessment communities and begin scoring student work jointly. Joint scoring could be used merely to establish and maintain common standards to enable teachers to assess their own students' work reliably, or it could be expanded to include "blind" scoring of student work so that each student receives feedback from another teacher.

With student work already online and identified only with unique identifiers, inter-institutional comparisons become possible. As long as faculty members who score student work maintain an inter-rater reliability correlation of 0.80 or higher, their scoring of student work can be used to compare learning outcomes among programs. Every student's work need not be scored because samples from each institution would be sufficient to generate reliable

... the process of shared portfolio assessment has generated discussion about what is valued in student writing, and participation in such discussions has increased faculty knowledge about writing and grading (R. Reagan, August 10, 2009).

data about learning outcomes for the purpose of reporting to the public. The process of linking local communities is similar to that of the VALUE system developed by the AAC&U, in which many institutions utilize the same rubric to score student work.

It is essential, of course, that the general abilities of students entering each institution are roughly comparable. This is why institutions should be able to choose their own peers for joint assessments. Each institution (or program) would be motivated to join a group that most closely matches its own level, neither significantly more nor significantly less difficult, similar to the formation of athletic conferences.

Eventually, as small groups of assessment communities merge into larger and larger groups, the process would work toward a set of national standards. Rather than a single standard, which would not account for the great diversity of missions and students served, several sets of standards could supply some latitude. The standards of liberal arts colleges that emphasize personal growth, for example, may be somewhat different from those of technical schools. The standards for all, however, would be transparent to the public, operationalized with examples and analysis, and available to students in the process of selecting schools. Reports to the public would consist of the percentage of the institution's students who fall into each scorepoint.

Both accrediting agencies and disciplinary associations could play an important role in this process if they chose to do so. The disciplinary associations have established memberships and rich sources of expertise and therefore can produce the most comprehensive, nuanced, and up-to-date assessments. Likewise, the accrediting agencies' knowledge of institutions in their purview could be used to help institutions identify peer groups.

Eventually, faculties' use of holistic scoring to assess student work and the gradual merger of local communities into larger groups could provide the framework for a national assessment. Validity would arise from faculties' initial—and continuing—use of the online assessment system. Reliability would be obtained through the continuing practice of assessment on the web site, so that teachers across the nation could refresh their assessment skills and compare their students' abilities with those exemplified in the national standards. Again, the operationalization of the standards through detailed analysis of actual samples of student work is fundamental.

Strengths and Limitations of This Method

Strengths.

- Online assessment communities are designed to achieve both high construct validity and high reliability.
- The in-depth analysis and discussion needed to produce high rates of inter-rater reliability result in powerful collaborative learning among participating faculty. The focus on faculty dialog reflects the value placed on faculty engagement in the Australian, European, and conventional American accreditation processes but even more centrally locates that dialog where it produces collaborative learning.
- Assessment results like tests scores, bar graphs, and other quantitative data are not readily usable to improve teaching, learning, and curricula. Online assessment communities, on the other hand, can deliver the kind of faculty development that can immediately affect teaching

Eventually, as small groups of assessment communities merge into larger and larger groups, the process would work toward a set of national standards.

practice and help “close the loop.”

Limitations.

- Like portfolio assessment and the use of rubrics like CLAQWA, online assessment communities require a commitment of time and effort from faculty. This “limitation,” however, is partly offset by the flexibility and convenience with which faculty can conduct their work online and by the meaningfulness and applicability of the assessment to the work of teaching.

Conclusion

Those of us in the academy must become far more deeply engaged in the assessment for accountability debate and far more articulate and forceful in explaining why authentic assessment is both efficient and effective in assessing student learning to demonstrate institutional accountability. Faculty are using authentic measures in their teaching and grading processes every day and these measures can do double duty in demonstrating accountability if aggregated at classroom, disciplinary, and institutional levels. We must develop and promote more authentic and valid assessments that are capable of accounting for the full range of learning outcomes rather than merely providing incomplete snapshots of basic skills. This is critical, because when assessment results are used to compare institutions, pressures on faculty to raise the publicly reported scores based on such snapshots will force us to spend more time, energy, and financial resources on learning activities that will increase those scores and to slight the far more difficult and demanding intellectual skills that are the true aim of a college education.

The methods described in this paper—electronic portfolios, common analytic rubrics, and online assessment communities—will not be quickly or easily scaled to a level that permits institutional comparisons. Although this will take a significant amount of faculty time, it will not consume more student time because assignments will simply be assessed twice—once for a grade and then again for accountability purposes. The motivation of students to do their best work is thus more likely than when assessment takes place on a single occasion.

If governments and funding agencies were to direct substantial funds to support faculty work on electronic portfolios, common analytic rubrics, and assessment communities, we could develop much richer forms of assessment for accountability purposes, yielding far deeper understanding of individual student learning and enabling individual students to learn far more about their own strengths and weaknesses. Also, by aggregating individual performances at department, college, and university levels, we can provide external stakeholders with extensive knowledge about the complexities of learning at individual institutions.

Authentic assessment of student learning outcomes is the way forward for improving pedagogy and programming as well as for demonstrating accountability. Let’s get started to develop these proven, promising methods on a much larger scale.

We must develop and promote more authentic and valid assessments that are capable of accounting for the full range of learning outcomes rather than merely providing incomplete snapshots of basic skills.

References

- American Philosophical Association. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. Millbrae, CA: Author. (ERIC Document Reproduction No. ED 315 423).
- Australian Universities Quality Agency (AUQA). (2009). *Setting and monitoring academic standards for Australian higher education: A discussion paper*. Melbourne, Australia: Author.
- Banta, T. W. (1999). Introduction: Why portfolios? In T. W. Banta (Ed.), *Portfolio assessment: Uses, cases, scoring, and impact* (pp. 1-5). San Francisco, CA: Jossey-Bass.
- Banta, T. W. (2007, January 26). A warning on measuring learning outcomes. *Inside Higher Education*.
- Banta, T. W., Jones, E. A., & Black, K. E. (2009). *Designing effective assessment: Principles and profiles of good practice*. San Francisco, CA: Jossey-Bass.
- Barr, R. B., & Tagg, J. (1995, November/December). From teaching to learning—A new paradigm for undergraduate education. *Change*, 27(6), 12-25.
- Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives: The classification of educational goals, Handbook 1. Cognitive domain*. New York, NY: McKay.
- Bruffee, K. A. (1999). *Collaborative learning: Higher education, interdependence, and the authority of knowledge*. Baltimore, MD: The Johns Hopkins University Press.
- Cambridge, D. (2009). Two faces of integrative learning online. In D. Cambridge, B. Cambridge, & K. B. Yancey (Eds.), *Electronic portfolios 2.0: Emergent research on implementation and impact* (pp. 41-49). Sterling, VA: Stylus Publishing.
- Day, M. (2009). Influencing learning through faculty- and student-generated outcome assessment. In D. Cambridge, B. Cambridge, & K. B. Yancey (Eds.), *Electronic portfolios 2.0: Emergent research on implementation and impact* (pp. 83-90). Sterling, VA: Stylus Publishing.
- Edwards, T. S., & Burnham, C. (2009). The promise of eportfolios for institutional assessment. In D. Cambridge, B. Cambridge, & K. B. Yancey (Eds.), *Electronic portfolios 2.0: Emergent research on implementation and impact* (pp. 87-90). Sterling, VA: Stylus Publishing.
- Elliot, N. (2003). *On a scale: A social history of writing assessment in America*. New York, NY: Peter Lang.
- Eynon, B. (2009). Making connections: The LaGuardia eportfolio. In D. Cambridge, B. Cambridge, & K. B. Yancey (Eds.), *Electronic portfolios 2.0: Emergent research on implementation and impact* (pp. 59-68). Sterling, VA: Stylus Publishing.
- Flateby, T. L., & Fehr, R. (2008). Assessing and improving the engineering curriculum. *International Journal of Engineering Education*, 24(5), 901-905.
- Flateby, T. L., & Metzger, E. A. (1999). Writing assessment instrument for higher order thinking skills. *Assessment Update*, 1999(March-April), 6-7.
- Flateby, T. L., & Metzger, E. A. (2001). Instructional implications of the cognitive level and quality writing assessment (CLAQWA). *Assessment Update*, 2001(January-February), 4-11.
- Griffin, M. (2009). Bridging the gap between college and high school writing in an online assessment community. *Dissertation Abstracts International*, 70-10A, (UMI No. 3381255).
- Hamilton, S., & Kahn, S. (2009). Demonstrating intellectual growth and development: The IUPUI ePort. In D. Cambridge, B. Cambridge, & K. B. Yancey (Eds.), *Electronic portfolios 2.0: Emergent research on implementation and impact* (pp. 91-96). Sterling, VA: Stylus Publishing.
- Huber, M. T., Hutchings, P., Gale, R., Miller, R., & Breen, M. (2007). Leading initiatives for integrative learning. *Liberal Education*, 93(2), 46-51.
- Kuh, G. D. (2008). Why integration and engagement are essential to effective educational practice in the twenty-first century. *Peer Review*, 10(4), 27-28.

- Kuh, G. D., Kinzie, J., Schuh, J. H., Whitt, E. J., & Associates. (2005). *Student success in college: Creating conditions that matter*. San Francisco, CA: Jossey-Bass.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York, NY: Cambridge University Press.
- Lewis, T., & Fournier, J. (2009). A catalyst without a mandate: Building an eportfolio culture at the University of Washington. In D. Cambridge, B. Cambridge, & K. B. Yancey (Eds.), *Electronic portfolios 2.0: Emergent research on implementation and impact* (pp. 129-132). Sterling, VA: Stylus Publishing.
- Lorenzo, G., & Ittelson, J. (2005). An overview of e-portfolios. Boulder, CO: EDUCAUSE Learning Initiative. Retrieved from www.educause.edu/ir/library/pdf/ELI3001.pdf
- Mintzberg, H. (2000). The professional bureaucracy. In M. C. Brown (Ed.), *Organization and governance in higher education* (5th ed., pp. 50-70). Boston, MA: Pearson.
- Reese, M., & Levy, R. (2009). *Assessing the future: E-portfolio trends, uses, and options in higher education* (Research Bulletin, Issue 4). Boulder, CO: EDUCAUSE Center for Applied Research.
- Schneider, C. G. (2008, Fall). From the president. *Peer Review*, 10(4), 3.
- Suskie, L. (2009). *Assessing student learning: A common sense guide* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Tagg, J. (2003). *The learning paradigm college*. Bolton, MA: Anker.
- Watson, C. E. (2007). *OSP 2.2—ePortfolio design and planning document*. Unpublished manuscript, Virginia Tech, Blacksburg, VA.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. New York, NY: Cambridge University Press.
- White, E. M. (1994). *Teaching and assessing writing* (2nd ed.). Portland, ME: Calendar Island Publishers.
- Yancey, K. B. (2009). Reflection and electronic portfolios: Inventing the self and reinventing the university. In D. Cambridge, B. Cambridge, & K. B. Yancey (Eds.), *Electronic portfolios 2.0: Emergent research on implementation and impact* (pp. 5-16). Sterling, VA: Stylus Publishing.

NILOA NATIONAL ADVISORY PANEL

Trudy W. Banta

*Professor
Indiana University-Purdue University
Indianapolis*

Douglas C. Bennett

*President
Earlham College*

Robert M. Berdahl

*President
Association of American Universities*

Molly Corbett Broad

*President
American Council on Education*

Judith Eaton

*President
Council for Higher Education Accreditation*

Richard Ekman

*President
Council of Independent Colleges*

Joni Finney

*Practice Professor
University of Pennsylvania*

*Vice President, National Center for Public
Policy and Higher Education*

Susan Johnston

*Executive Vice President
Association of Governing Boards*

Paul Lingenfelter

*President
State Higher Education Executive Officers*

George Mehaffy

*Vice President
Academic Leadership and Change
American Association of State Colleges and
Universities*

Margaret Miller

*Professor
University of Virginia*

Charlene Nunley

*Program Director
Doctoral Program in Community College
Policy and Administration
University of Maryland University College*

Randy Swing

*Executive Director
Association for Institutional Research*

Carol Geary Schneider

*President
Association of American Colleges and
Universities*

David Shulenburg

*Vice President
Association of Public and Land-Grant
Universities*

Belle Wheelan

*President
Southern Association of Colleges and Schools*

George Wright

*President
Prairie View A&M University*

Ex-Officio Members

Peter Ewell

*Vice President
National Center for Higher Education
Management Systems*

Stanley Ikenberry

*Interim President
University of Illinois*

George Kuh

*Chancellor's Professor
Indiana University*

NATIONAL INSTITUTE FOR LEARNING OUTCOMES ASSESSMENT (NILOA) MISSION

NILOA's primary objective is to discover and disseminate ways that academic programs and institutions can productively use assessment data internally to inform and strengthen undergraduate education, and externally to communicate with policy makers, families and other stakeholders.

NILOA OCCASIONAL PAPER SERIES

NILOA Occasional Papers are commissioned to examine contemporary issues that will inform the academic community of the current state-of-the art of assessing learning outcomes in American higher education. The authors are asked to write for a general audience in order to provide comprehensive, accurate information about how institutions and other organizations can become more proficient at assessing and reporting student learning outcomes for the purposes of improving student learning and responsibly fulfilling expectations for transparency and accountability to policy makers and other external audiences.

Comments and questions about this paper should be sent to sprovez2@illinois.edu.



ABOUT NILOA

- The National Institute for Learning Outcomes Assessment (NILOA) was established in December 2008.
- NILOA is co-located at the University of Illinois and Indiana University.
- The NILOA web site went live on February 11, 2009.
www.learningoutcomesassessment.org
- The NILOA research team reviewed 725 institution web sites for learning outcomes assessment transparency from March 2009 to August 2009.
- One of the co-principal NILOA investigators, George Kuh, founded the National Survey for Student Engagement (NSSE).
- The other co-principal investigator for NILOA, Stanley Ikenberry, was president of the University of Illinois from 1979 to 1995 and of the American Council of Education from 1996 to 2001. He is currently serving as Interim President of the University of Illinois.
- Peter Ewell joined NILOA as a senior scholar in November 2009.

NILOA STAFF

NATIONAL INSTITUTE FOR LEARNING OUTCOMES ASSESSMENT

Stanley Ikenberry, *Co-Principal Investigator*

George Kuh, *Co-Principal Investigator and Director*

Peter Ewell, *Senior Scholar*

Staci Provezis, *Project Manager*

Jillian Kinzie, *Associate Research Scientist*

Jason Goldfarb, *Research Analyst*

Natasha Jankowski, *Research Analyst*

Gloria Jea, *Research Analyst*

Julia Makela, *Research Analyst*

NILOA SPONSORS

Carnegie Corporation of New York

Lumina Foundation for Education

The Teagle Foundation

educate innovation success ingenuity intellect curiosity challenge create achievement knowledge accountability connection self-reflection educate a
nderstand communicate curiosity challenge create achievement connection self-reflection understand communicate listen learn access quality acti
ction understand communicate listen learn action understand communicate listen learn access quality innovation success ingenuity intellect curios
challenge knowledge accountability connection access quality self-reflection curiosity challenge create achievement learn access quality innova
ingenuity self-reflection educate action understand intellect knowledge accountability connection self-reflection educate action understand knowled
accountability connection self-reflection educate action understand communicate listen learn access quality innovation success ingenuity intellect c
hallenge connection knowledge accountability connection self-reflection educate action understand communicate listen learn access quality innova
uccess ingenuity challenge create achievement connection self-reflection educate action understand connection self-reflection understand commun
sten learn access quality action create achievement connection self-reflection educate action understand communicate listen learn access quality
uccess educate action communicate listen learn access quality action educate action understand communicate educate innovation success self-ref
knowledge accountability communicate listen learn achievement connection self-reflection educate action understand communicate listen learn acc
novation success ingenuity intellect access quality innovation success self-reflection curiosity challenge create achievement connection self-reflec
nderstand educate action understand communicate listen learn action understand communicate listen learn access quality innovation success ing
uriosity challenge create achievement connection self-reflection understand communicate listen learn access quality action create achievement co
self-reflection educate action understand communicate listen learn access quality innovation success educate action communicate listen learn acc
pality action educate action understand create achievement connection self-reflection understand communicate listen learn access quality action c
chievement connection self-reflection educate action understand communicate listen communicate educate innovation success self-reflection know
accountability connection self-reflection educate action understand communicate listen learn access quality innovation ingenuity intellect connectio
reflection understand communicate listen learn access quality action create achievement connection self-reflection educate action understand comm
sten learn access quality innovation success educate action communicate listen learn access quality action educate action understand communica

NATIONAL INSTITUTE FOR LEARNING OUTCOMES ASSESSMENT

FOR MORE INFORMATION, PLEASE CONTACT:

National Institute for Learning Outcomes Assessment (NILOA)
University of Illinois at Urbana-Champaign
340 Education Building
Champaign, IL 61820

learningoutcomesassessment.org
sprovez2@illinois.edu
Fax: 217.244.3378
Phone: 217.244.2155