

A DECADE OF RESEARCH: ASSESSING CHANGE IN THE TECHNICAL COMMUNICATION CLASSROOM USING ONLINE PORTFOLIOS

CAROL SIRI JOHNSON

New Jersey Institute of Technology

ABSTRACT

Over a period of 10 years, we have developed a sustainable process of online portfolio assessment that demonstrates both reliability and validity, using both qualitative and quantitative measures. The sustainable cycle is that, each semester, we assess a random sampling of the students' work that they have posted, as per our instructions, in an online portfolio. During the reading, the faculty score the documents for 11 variables, including writing, content, audience awareness, and document design. We achieved validity by a modified online Delphi that led to a redefinition of the construct of technical communication itself; we achieved reliability by adjudication resulting in adjacent scores. The results of our assessment meet the requirements of ABET and result in a continual cycle of improvement for our technical communication curriculum. Results from three semesters show an improving correlation between the course grade and the overall, holistic portfolio score.

CHANGING VIEWS OF ASSESSMENT

In *(Re)Articulating Writing Assessment*, Brian Hout writes "Because assessment is the site where we marshal evidence about what we will value globally as a society and more locally as teachers, researchers and administrators, we can, by changing assessment, change what we will ultimately value" [1]. Paradigms of writing assessment have changed considerably within the past 10 years. Holistic best-paper readings, developed for gate keeping and placement purposes, are being replaced by new forms of assessment that seek greater construct validity and a

wider range of inputs [2]. The Accreditation Board of Engineering and Technology (ABET) has also mandated that each program use multiple assessment measures and, further, that the outcomes be used to improve the program [3]. The argument of this article is that we should, if we can, incorporate a culture of assessment into our programs, designed by ourselves, conducted by ourselves, informed by the history of educational and writing assessment and resulting in a sustainable process of self evaluation and improvement. Huot agrees with these arguments. He has pointed out that “writing assessment must be site-based and locally controlled” and that “writing teachers and program administrators must begin to see writing assessment as part of their jobs.” In this article, I will describe the results of a 10-year effort to incorporate a sustainable forum for self-assessment in the technical communication service course. We have achieved both sustainability and the requirements of educational assessment—reliability, validity—by assessing online portfolios.

During most of the 20th century, validity (the match between the test and its purpose) was sacrificed in order to gain reliability (stability in scoring) [4]. One goal of assessment in the 21st century is to gain more construct validity by bringing theory, evidence, and purpose into balance. The search for validity itself is important. According to the *Standards for Educational and Psychological Testing*, validity is “the most fundamental consideration in developing and evaluating tests” [5]. According to the *Standards*, this includes “delineating the knowledge, skills, abilities, processes, or characteristics to be assessed.” Thus, our goal is to achieve greater construct representation and balance it with a careful sampling plan that can capture the qualities inherent in the construct itself (rather than in a single result). Our assessment process led us to redefine our construct of technical communication and we incorporated that revised construct back into the assessment. In the 21st century, assessment is becoming more multifaceted, cyclical and recursive.

Writing is a complex activity that enables in-depth learning in a wide variety of fields. As Roberta Camp of the Educational Testing Service has observed, writing is “a rich, multifaceted, meaning-making activity that occurs over time and in a social context, an activity that varies with purpose, situation, and audience and is improved by reflection on the written product and on the strategies used in creating it” [6]. In her longitudinal study of college writers, Marilyn S. Sternglass observed that instruction in writing is critical to fostering an analytic stance in disciplinary courses, instruction that allows students to “understand the significance of ideas in their particular field to the level where they become able to question some of the assumptions of that field” [7]. In the field of engineering education, Carolyn Miller has praised the communication-across-the-curriculum movement in her review of program outcomes established by ABET. That movement, she wrote, “aims to incorporate writing and speaking into every program and department across our campuses in order to demonstrate that these are not isolated subjects but rather arts that are relevant to every subject” [8].

Thus, in the 21st century, writing is far more than a skill: it is a vehicle for empowerment [9].

Even as we seek to empower students, we can empower ourselves by taking responsibility and control of our assessment processes. One feature in the evolution of assessment is that it is increasingly created by teachers for teachers. Although there must be a guide for the process, that guide can seek input from faculty, instructors, professionals in the field, and students, and thus combine the best of ideas from many people to create a culture of consensus. This change has been documented in engineering institutions. In a recent study, “Quality Assurance of Engineering Education through Accreditation: The Impact of Engineering Criteria 2000 and Its Global Influence,” John Prados, George Peterson and Lisa Lattuca have noted that more departmental members are taking part in assessment processes than ever before. Their first two bullet points regarding the faculty involvement in the ABET processes are:

- 70% of program chairs report high levels of faculty support for continuous improvement efforts; and
- 88% of faculty members report at least some personal effort in program assessment [10].

Thus, ABET is encouraging greater participation in assessment. We, as technical writing teachers, can use this encouragement to take control (authority) of assessment in our departments and make a sustainable culture of participatory inquiry resulting in continual evaluation, change, and improvement. ABET’s EC2000 is calling for the same sort of site-specific, contextually-driven assessment practices that the field of composition seeks. As teachers in the field of technical communication, we are uniquely placed as a crossover between composition and engineering; we can take this opportunity to create a new way of designing, implementing, and incorporating grassroots assessment in our curriculum. The support of the department chair and college dean is necessary in order to gain practical access to the conference rooms, copying, coffee and other physical equipment required. Otherwise, the materials for self-assessment are already in our hands.

BACKGROUND: A CULTURE OF ASSESSMENT

Our university is one of the most diverse in the nation. Although the statistics do not tell the whole story, undergraduate enrollment in 2004 was 10.5% African American, 21% Asian, 12.6% Hispanic, 34% White, 6% non-United States, and 15.4% unknown. Also, the number of U.S. residents who speak English as a second language is approximately half [11]. Thus, we must communicate across language boundaries that mirror a global setting. This reality has led us to increased innovation and to include more visual training in the technical communication classroom.

The Department of Humanities has a long history of assessment research and practice. Robert Lynch, the Department Chair, was instrumental in the development of the New Jersey Basic Skills test, administered to all incoming freshmen at state institutions, and in the early development of the ABET criteria. Norbert Elliot has just published *On a Scale: A Social History of Writing Assessment in America* [12]. Most of our faculty take part in assessment every semester. Each of our GURs (General University Requirements) has its own form of program assessment—Freshman Composition has best paper readings, Cultural Studies and the Capstones have best paper and portfolio readings, and the Master's in Professional and Technical Communication has both a formative and summative assessment of online portfolios. We have a Website on College Writing Assessment which includes resources at <http://cwa.njit.edu/>. This culture of assessment ebbs and flows as the need arises: we use assessment as a focused tool for inquiry, to explore and improve our courses.

The 10-year project of assessing technical communication portfolios was begun by Norbert Elliot, Margaret Kilduff, and Robert Lynch in 1988, and is described in “the Assessment of Technical Writing: A Case Study” [13]. This article describes “a first-generation effort to come to terms with issues and methodologies of writing assessment that are appropriate to the technical writing class.” Elliot, Lynch, and Kilduff first sought input from others to create a set of interrelated writing tasks. The next step was to provide a method for assessing the performance of students related to those tasks. In order to do so, they created a rubric on a 4-point scale to indicate levels of competence on the presentation of ideas, cohesion, style, usage, and overall reader response. The portfolios were scored by two readers and the discrepant readings were adjudicated by a third. Thus, they achieved both reliability and validity.

However, collecting paper portfolios with multiple drafts of multiple documents proved too cumbersome to allow for the assessment of entire portfolios, so they assessed clusters of best samples. In the three years of this study (1989-1991), 308 clusters were evaluated. Even so, the sheer physical volume of paper portfolios was a problem and, thus, this portfolio assessment method was not sustainable. Nevertheless, it initiated a process, according to Elliot, that “creates consensus among instructors and yields valuable information to students, faculty, and administrators alike.” The aspects of consensus and community were significant enough to take hold and provide the basis for further years of research. Indeed, they may be the most important aspects of cyclical programmatic assessment. Nancy Coppola continued the inquiry in “Setting the Discourse Community: Tasks and Assessment for the New Technical Communication Service Course,” published in 1999 [14]. Coppola notes that reading portfolios “increases collaboration between teachers and program directors by bringing the course content into a public setting.” Thus, she views “portfolio assessment [as a] social construction in practice.” The very act of assessment, then, creates a sense of

community and knowledge is created through interactions, from student to student, from student to instructor and from instructor to student and between instructors as well.

ASSESSMENT AND IMPROVEMENT

For several years after the Elliot study, the program developers went on to other administrative posts and new leadership in the Department of Humanities discontinued the cycles of assessment. Therefore, instructors taught independently, choosing their own texts, writing their own syllabi, and rarely communicating with each other; technical communication had truly become a service course, running on automatic. In the spring of 2003, leadership changed again and we resumed a departmental assessment effort for all of the GURs, using holistic best paper readings as a forum for curricular discussion and improvement. At that time I became the coordinator of the technical communication course and began to learn how to use the elements of assessment. On reading day, instructors—untentured and tentured, one-year appointments and adjuncts—gathered together at the end of the semester and, after a calibration session with sample papers, used an old rubric (written for research papers) to holistically score the papers on a 6-point scale.

With this first assessment after the hiatus, using best papers and an old rubric, our mean was 7.396 (on a scale from 2 to 12), which told us that we were basically reaching our goals. More importantly, however, this assessment led to more interaction and communication within the department. First we discussed and chose a common text to use for the fall of 2003. Due to this simple factor—renewed assessment of our program—the mean scores in the fall immediately rose from 7.396 to 8.328, a significant improvement ($t = 2.4828$; $p = .0072$). The next step was to bring the course into line with methods of reading and writing technical communication in the 21st century: during the time that Elliot began the study, PCs had replaced paper and students were literate in a new way. We wanted to use the outcomes assessment to look for trend information and patterns in student learning that could inform curricular change. In the spring of 2004, after much discussion with professors, instructors, and professional technical writers from the Society for Technical Communication, we created a series of interrelated writing projects that represented our construct of technical communication and instituted them as required modules within the larger course.

The common modules were based on modes of discourse from both industry and academics: we required a procedure, a proposal, a marketing brochure, an oral presentation, and a Website. The procedure and proposal were the anchors of the design; the brochure and Website were new visual elements that would require different forms of thinking and creating documents. The five common modules were:

- **Manual** *Task:* Write a technical procedure, or set of steps, for a specific action, such as how to use a specific function in Excel, PhotoShop, or AutoCAD, how to use a machine, how to modify an automobile, etc.
Result: MS Word document, at least five pages long with headings, numbered steps, graphics, and a table of contents.
Audience: A moderately skilled user.
- **Proposal** *Task:* Write a proposal including the following sections (if necessary): cover letter, executive summary, table of contents, background, objectives, plan, personnel, budget, time line, conclusion, supporting materials.
Result: MS Word document, at least five pages long, with formal and consistent writing, grammar, and mechanics (no errors or typos).
Audience: Decision-makers within the organization.
- **Technical Marketing Brochure** *Task:* Use graphic and text programs to create an attractive marketing brochure for the topic in your proposal.
Result: A two-sided brochure, in color, incorporating both graphics and text in a carefully crafted layout.
Audience: Potential purchasers.
- **Oral Presentation** *Task:* Prepare an individual presentation that will be judged by other class members for clarity, interest, etc. on a presentation report.
Result: You will receive written assessments from each classmate.
Audience: Class members.
- **Website** *Task:* Create a Website with links to the manual, proposal, and brochure. The documents can be in Word, PDF, or HTML.
Result: URL for presentation to the class and for the final assessment.
Audience: Technical communication professors and classmates.

This set of common elements ensured that the class would work through five separate modes of discourse—instructional, persuasive, visual, oral, and online. The instructors were free to decide how to weigh each assignment for grading and to add their own assignments as well. The benefit of this set of assignments was that it would lead the core activities during the semester and result in a paperless portfolio that we could easily assess online.

We began using these new modules in the spring of 2004. This was a large change for many of the instructors since most of them had never taught Web design. Moreover, creating the online portfolio at NJIT was complex: students had to sign up for Web space, configure the space for Web hosting, and then create

and upload a Website. Each of these steps could be fraught with problems and errors. I provided a basic HTML template and set of instructions for building the Website and, surprisingly, even these tools were unnecessary for some of the students—they knew how to use many different tools to create their online portfolios. We then grouped students into teams so that they could teach these methods to each other. In this case, the role of the instructor changed from teacher to facilitator.

The basic required Website was simple: at the least, it needed to have each student's name and links to their procedure, proposal, and brochure. The template that I provided had these simple features written in basic HTML with highlighted text and links so that they could replace it with text and links of their own. Figure 1 is an example of a basic portfolio because it follows the instructions exactly; the author offers no more, no less.

Of course, many students went beyond the basic template and added art, personal material, and links to their other work. Many of these portfolios evolved into semi-permanent repositories of the student's work and life. Figure 2 illustrates an advanced portfolio that has continued to grow during the student's time at NJIT, with the addition of links to resumes and other courses, multimedia, and

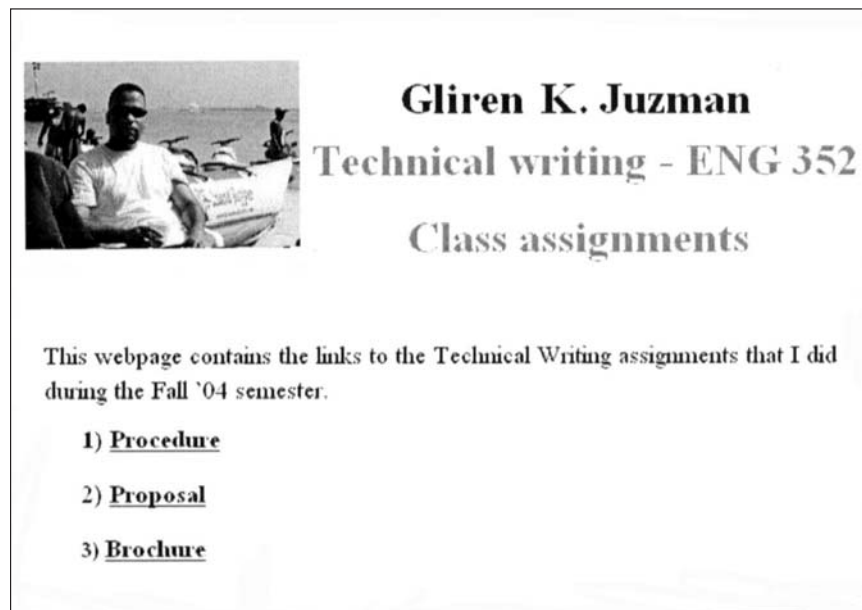


Figure 1. Basic online portfolio created by using a simple HTML template.



Figure 2. Advanced online portfolio with multimedia, multiple paths, and modes of discourse.

art. It is also more complex in that it has multiple modes within each page and multiple pathways between the pages, resulting in a multi-dimensional hypertext environment.

In “Postmodernism, Palimpsest, and Portfolios: Theoretical Issues in the Representation of Student Work,” Kathleen Blake Yancey describes a “Web sensible” portfolio that has multiple navigation paths and makes full use of digital media [15]. She notes that an online portfolio has “a textured literacy that is different in kind than the thesis-and-support literacy of the print model” and compares this sort of portfolio to a gallery. One of the great strengths of the individually-composed online portfolio for students is that they *are* a gallery—the students are using all aspects of their mind—analytic, visual, auditory, and verbal—to show us who they are. Thus, they are able to learn more rapidly from multiple sources and they truly care about the results and, thus, according to Yancey, “[portfolios] perform a double function—providing grist for the twin mills of identity and assessment.”

It is important that students have a stake in the assessment, even if it is only self-expression. Yancy also describes a closed “online assessment system” where students upload their work to a remote database, often created by an outside

vendor, in order to provide writing for the purposes of ongoing assessment. This system offers more administrative control but it removes control from the students. The students are no longer stakeholders and thus they view submitting their writing to a closed database as a frustrating and pointless chore. In order to make assessment sustainable, all parties—faculty, instructors, administration, and students—must have an active role and an interest in the outcome.

What We Really Value: Developing New Criteria

In the spring of 2004 we switched from reading best papers to reading the online portfolios using an old portfolio score sheet developed for the GURs, which were mainly Humanities Courses. This first attempt to read online portfolios failed miserably because of a technical problem—we were using wireless PCs during a time when the network was experiencing viruses. Thus, we took the opportunity to discuss the old portfolio score sheet and how to improve it. In a way, the virus was a lucky accident because we needed to step back and assess our own procedures and the validity of the rubric. In *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing*, Bob Broad notes that, in writing assessment, we achieved reliability at the cost of validity [16]. He advocates “Dynamic Criteria Mapping,” an assessment process in which he linguistically analyzes the verbal interactions between writing teachers as they are scoring student writing. His underlying assumption is that “people do not have satisfactory access to their rhetorical values by sitting and reflecting on them. Instead, people need to enter into a discussion and debate of actual performances in an effort to discover what they (and others) value.” Thus, we asked ourselves, “What do we assess?” After recording our discussion for eventual analysis, we extended the conversation through the use of e-mail.

We began with a list of core competencies compiled by Nancy Coppola and Norbert Elliot for the Master’s of Professional and Technical Communication Program (MSPTC), described in “Assessment of Graduate Programs in Technical Communication: Constructing a Model to Engage Complexity,” an article forthcoming in *Assessment in Professional and Technical Communication*, edited by Jo Allen and Margaret Hundley [17]. Coppola and Elliot used contemporary technical communication research to construct an assessment matrix of core competencies. This matrix was based on the data reported by David Dayton and Stephen A. Bernhardt in “Results of a Survey of ATTW Members, 2003,” a study in which they collected information from members [18]. One query asked respondents to rate the importance of certain skills to professional life. The results were, in order of importance: rhetoric, writing and editing, technology, personal traits and work skills, specialized expertise, document design, problem solving/thinking skills, collaboration and teamwork, oral or interpersonal communication, and research. The MSPTC matrix used these as headings for the core competencies which were then broken down further into descriptors.

For the technical communication service course, we needed to narrow the focus to a one-semester undergraduate educational experience. After making a basic list of criteria, we held a series of modified online Delphi sessions to gain input and consensus on the changed course goals. According to Harold Linstone and Murray Turoff in *the Delphi Method: Techniques and Applications*, a Delphi can be defined as “a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem” [19]. It can also be used, as we did, as a way of querying and resolving policy questions. After the initial discussion with the instructors of English 352, we moved the discussion to professors of technical communication. The discussion took place in four e-mail threads, which were mainly linear but sometimes branched off into individual discussions. The following is an exact transcription of the main thread in this discussion, which took place over two weeks. This discussion represents change in the teaching of technical communication—in the past, there would have been no discussion. The view of what technical communication consisted of was received knowledge. Herman Estrin, one of the founding fathers of technical writing in the United States, taught technical communication as a didactic and rule-driven process. Headings, layout, and processes were all prescribed. However, in 21st century assessment, one of our tasks is a search for language, a new terminology to discuss, and evaluate the changing world. The following discussion is a negotiation of terminology.

Moderator (in original e-mail): Since I have been discussing this and reading about it, I have come across many terms to refer to it. It used to be called a rubric. In the MSPTC version, it is a list of “Core Competencies” divided further into “Descriptors” and “NJIT Descriptors”; the entire set is also called a “Matrix.” Broadhead and Freed call them “Variables” and Bob Broad calls them “Criteria.” In these e-mails, I have been calling them “New Criteria.” If I am going to publish about this, I want to use a name that is easily understandable to the largest number of people. I also want to use it consistently. Should I continue to use “New Criteria” or should I use another name? Any ideas?

Responder A: I suppose a new terminology sets you and your work apart somehow, and allows you to look at things slightly differently. But is there a need for a new terminology?

Responder B: I would think that variables are good in that they connote association. We want to think about the aspects of technical writing that are associated with success. Criteria always sounded like a bar that was being set in track meet. So, it is a matter of this: Do we want a behavioral context or a standard-bearing context? It is all in the language, of course. We all know how I feel about avoiding bars in all their aspects—avoid, at all costs.

Moderator: What I really want is a common terminology. Responder B suggested that we refer to these criteria as variables. do we all agree that the

term “variable” can express the different qualities we want to find in technical writing?

Responder C: I don’t see what’s wrong with “criteria,” which carries a sense of fixedness, and means standards and rules used to base a determination (like, on what basis do we judge a text as being adequate for its purpose?). “Variable” is something that varies or is prone to variation, or in math, a quantity capable of assuming any value. If you’re looking for a word that comes close to “qualities,” I think you’re better off with criteria than with variable.

Responder B: It is always a series of contingencies in technical writing. The more we educate our students to create audiences and to understand uncertainty, the better prepared they will be. In addition, I mean variable in terms of relation to an outcome—dependent variables lead to independent variables—just a bit more subtle definition, I think.

Responder D: I think the word choice depends on the context (which is why, for one reason, we used different words to express similar meanings in the graduate program study.) “Variables” has a specific meaning in a research context; “criteria” is a broader, more accessible word. We chose competencies because we wanted to stress what the student/graduate should be able to do after enrollment in our program; we wanted to look at the outcomes rather than the measurements of the success. I think you have to decide what you are really trying to capture and, thus, define the context.

Responder A: As possibly a non-expert in the field of technical communication, ultimately I will nonetheless agree with Responder C here.

Moderator: The purpose of this 352 assessment Delphi is to define what is important to teach in a technical communication class for engineers. By naming these variables or criteria, we will create a basic structure for the course. The teaching of literacy has always been in flux—whether literacy is defined as knowledge of Latin, Greek, English grammar, or literature—and now, because of the information age, it is again changing. What do we teach? And why?

I have a glossary of assessment terms in front of me. They were made up for the WPA Assessment Institute I went to over the summer. Although the word “variable” never appears, the word “criteria” appears many times. Also, the first definition of “criteria” in the dictionary is “a standard of judgment or criticism; a rule of principle for evaluating or testing something” and definitions 1-9 for “variable” are about change—it isn’t until definition 10 that a variable is defined in mathematical terms. Thus, although these criteria are variable, to use the word variable would confuse (and exclude) some readers.

So 3 for criteria, 1 for variable. Could you, Responder D and Responder E, vote?

Responder E: Only in academia can an exchange like this involve so many people! I'll go with "criteria" so long as you use it as a plural. (The singular is "criterion.")

After this initial exchange of ideas, we held more online discussions and thus agreed on the major categories for the criteria and then on the criteria themselves. Only one item on our list, Textual Attribution, did not achieve consensus. After the online Delphi sessions, we met to discuss and debate what we had previously discussed in the e-mails. Although we still could not reach consensus on Textual Attribution, we were able, through discussion, to achieve consensus on a set of 11 new criteria for assessing undergraduate technical communication. The criteria (Figure 3) were then posted on the Web, distributed to the instructors, and used on the score sheets for the following assessment.

We could not achieve consensus on the two variables (or criteria) in "Textual Attribution" because they had a dual purpose—to check for proper attribution and also to check for plagiarism and originality. We were not able to agree on the wording for this and thus the scoring of those two items was optional.

CONDUCTING AN ANALYTIC ONLINE PORTFOLIO READING

In the fall of 2004, we had a successful online portfolio analytic reading (using hardwired computers to protect against viruses). We used a randomized sample of 88 students to reach a confidence interval of 8, each of which submitted a single sheet of paper with their name and URL. The English 352 instructors gathered for breakfast and a calibration session (recorded) and then moved to the hardwired lab to read and score the online portfolios. The scoring sheet was two pages, with the above criteria scored on a Likert scale. Each portfolio was read by two readers. If they were discrepant—if any criteria were scored more than one point apart—those criteria were highlighted on a new score sheet and given to a third reader to resolve the discrepancy. Due to the length and the complexity of the scoring procedure, we only managed to read 39 portfolios on that day. Later, URLs were distributed to the instructors to bring the sample size up to 60.

RESULTS

Reliability

For the null hypothesis significance testing (NHST), I used the Pearson correlation coefficient and found a significant correlation between readers in all variables except 8 (the contents of this portfolio demonstrate that the author can adapt design for purpose), which we then excluded from further testing (see Figure 4). Thus, we achieved a 95% confidence level in inter-reader reliability.

English 152, Technical Communication Portfolio Assessment Fall 2014						
Student Name	Score					
Teacher Name						
<p>Document Design The contents of this portfolio demonstrate cohesion by graphic means (headings, white space) in documents.</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p> <p>The contents of this portfolio demonstrate use of parallel structure with heading and sub-headings.</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p> <p>The contents of this portfolio demonstrate use of basic graphics.</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p>						
<p>Textual Attribution The contents of this portfolio demonstrate the ability to differentiate between sources and one's own text.</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p> <p>The contents of this portfolio demonstrate uniform textual citation.</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p>						
<p>Overall Portfolio Score</p> <table border="1"> <tr> <td>The materials in the portfolio demonstrate very good work in the class.</td> <td>The materials in the portfolio demonstrate average work in the class.</td> <td>The materials in the portfolio demonstrate below average work in the class.</td> <td>The materials in the portfolio demonstrate work that is at a level of one level of below average work in the class.</td> <td>The materials in the portfolio demonstrate work that is at a level of below average work in the class.</td> </tr> </table>		The materials in the portfolio demonstrate very good work in the class.	The materials in the portfolio demonstrate average work in the class.	The materials in the portfolio demonstrate below average work in the class.	The materials in the portfolio demonstrate work that is at a level of one level of below average work in the class.	The materials in the portfolio demonstrate work that is at a level of below average work in the class.
The materials in the portfolio demonstrate very good work in the class.	The materials in the portfolio demonstrate average work in the class.	The materials in the portfolio demonstrate below average work in the class.	The materials in the portfolio demonstrate work that is at a level of one level of below average work in the class.	The materials in the portfolio demonstrate work that is at a level of below average work in the class.		

English 152, Technical Communication Portfolio Assessment Fall 2014	
Student Name	Score
Teacher Name	
<p>Writing and Editing The contents of this portfolio exhibit clear style (readable, concise, coherent).</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p> <p>The contents of this portfolio demonstrate accurate language usage (grammar, punctuation, spelling).</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p>	
<p>Substance and Content The contents of this portfolio exhibit clear understanding of assignments.</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p> <p>The contents of this portfolio demonstrate accurate, thorough, relevant, and coherent content and ideas.</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p> <p>The contents of this portfolio demonstrate that the author can respond to different rhetorical situations.</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p>	
<p>Audience Awareness The contents of this portfolio demonstrate that the author can adapt content to audience and purpose.</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p> <p>The contents of this portfolio demonstrate that the author can adapt tone for audience.</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p> <p>The contents of this portfolio demonstrate that the author can adapt design for purpose.</p> <p>Very Strongly Agree Strongly Agree Agree Disagree Strongly Disagree Very Strongly Disagree</p>	

Figure 3. New criteria on score sheet.

Variable	Pearson coefficient
1	0.760349
2	0.811497
3	0.712113
4	0.747054
5	0.740192
6	0.695074
7	0.734047
8	0.57529
9	0.830157
10	0.786503
11	0.724889
14	0.770856

Figure 4. Pearson correlation coefficients for adjudicated scores in Fall 2004.

After disproving the null hypothesis, I calculated the means of each variable. This was where we hoped to find patterns. The mean score from all variables was 7.888 and the overall portfolio score was 7.869. These scores demonstrate that we have an effective faculty with a high level of performance and dedicated students who are learning the material. Figure 5 presents the adjudicated scores over three semesters. The reliability increased steadily from semester to semester: we were becoming more comfortable with the assessment process and more calibrated as a group.

Historically, reliability has been a major issue in assessment and, thus far, no one has scored online portfolios reliably. We conducted an analytic scoring, scoring for 11 separate criteria. We achieved reliability on all of those criteria but one. Thus, it can be done. It did not cost money but it did take time, energy, and departmental commitment.

Validity

Not only did we read reliably, but through the creation of the new criteria, we met the conditions of content and construct validity. Historically, most portfolios

Criteria	Descriptors	Inter-reader Reliability ¹						Overall Portfolio Scores					
		Adj. Pearson's r^2		Cronbach's Alpha		Mean		Standard Deviation					
		2004 ²	2005 ³	2004 ²	2005 ³	2004	2005	2004	2005	2004	2005		
Writing and Editing	Exhibits clear style (readable, concise, cohesive)	.449	.473	.821	.686	.792	.898	7.77	7.94	8.19	1.283	1.476	1.948
	Demonstrates accurate language usage (grammar, punctuation, spelling)	.564	.639	.876	.816	.808	.919	7.54	7.76	7.77	1.409	1.572	2.163
Substance and Content	Exhibits clear understanding of assignments	.276	.555	.860	.709	.773	.921	8.00	8.32	8.20	1.225	1.377	2.048
	Demonstrates accurate, thorough, relevant, and coherent content and ideas	.642	.509	.828	.747	.807	.887	7.93	7.84	8.25	1.195	1.530	1.885
	Can respond to different rhetorical situations	.308	.645	.839	.756	.779	.912	7.95	7.96	8.22	1.359	1.511	1.945
	Can adapt content to audience and purpose	.418	.603	.816	.695	.751	.907	8.02	7.90	8.29	1.359	1.542	1.920
Audience Awareness	Can adapt tone for audience	.468	.607	.801	.747	.785	.885	7.82	7.94	8.19	1.272	1.434	1.928
	Demonstrates cohesion by graphic means (headings, white space) in documents	.565	.616	.840	.769	.855	.912	7.84	8.00	8.10	1.529	1.666	1.977
Document Design	Uses parallel structure with heading and subheadings	.466	.579	.841	.769	.769	.907	7.59	8.08	8.05	1.407	1.455	1.950
	Includes basic graphics	.379	.456	.813	.664	.728	.891	8.34	8.82	8.56	1.250	1.273	1.791
	Overall Portfolio Score	.584	.707	.839	.782	.746	.912	7.82	8.08	8.45	1.397	1.441	1.997

¹ In all cases, the correlation is significant at the 0.01 level (2-tailed).
² Adjudicated scores include third readings for discrepancies.
³ Fall 2004 (n=61).
⁴ Spring 2005 (n=50).
⁵ Fall 2005 (n=136).

Figure 5. Inter-reader reliability for fall 2004 to fall 2005.

have been read by using a single score in a holistic reading. We articulated the values inherent in our course and in our teaching and created an instrument to reliably score those variables. Thus, the conditions of reliability and construct validity have been met. Moreover, we have data that we can query in a variety of ways and, as our data pool grows, we will be able to refine our methods to pose new questions.

Searching for Patterns

Our original goal was to see if online portfolios could be scored reliably and with construct validity. Since we achieved a stable method of online portfolio assessment, we began to search for patterns in the results. One of our concerns was that the grades, for the past several years, had been higher than those for other Humanities courses. Although we were achieving our goal of a minimum of 7 on the online portfolios, there was a disjuncture between the assessment scores and the course grade. In fact, in the fall of 2004 there was no significant correlation between the course grade and the overall portfolio score. However, the correlations grew during the next three semesters to become significant. Figure 6 shows the changes in correlations. This meant that our instructors were grading the students with more accuracy and greater attention to our articulated common goals.

CONCLUSION

In this article, we have shown that assessment can be integrated into an academic community so that all people involved can make choices and take part. Moreover, we have shown that online portfolios can be read analytically, scored with reliability and validity, and that the scores can be used for further inquiry. Assessment can and should be an integral part of every department and it should be designed and run by the department. However, since it is labor-intensive, the department should be prepared to support the effort, even if it means providing further education for its practitioners, release time for the extra work and an occasional breakfast or lunch.

This new form of portfolio assessment also fulfills the ABET EC2000 criteria—it uses assessment to continually improve programs. Ongoing cycles of assessment can provide a basis for collaboration and intellectual exchange that can help us to review and revise our own criteria, to look at ourselves and our programs critically, to make changes, and to query those changes. A ten-year process of research, discovery and collaboration went into creating this paradigm. It is valuable in that it uses both qualitative and quantitative measures and incorporates new theories of assessment with the established theories of reliability and validity. It is within our power to use assessment to help us adjust to change in a continually changing world.

Fall 2004		Overall Portfolio Score	Course Grade
Overall Portfolio Score	Pearson Correlation	1	.119
	Sig. (2-tailed)		.363
	N	61	61
Course Grade	Pearson Correlation	.119	1
	Sig. (2-tailed)	.363	
	N	61	61

Spring 2005		Overall Portfolio Score	Course Grade
Overall Portfolio Score	Pearson Correlation	1	.399(**)
	Sig. (2-tailed)		.004
	N	50	50
Course Grade	Pearson Correlation	.399(**)	1
	Sig. (2-tailed)	.004	
	N	50	50

** Correlation is significant at the 0.01 level (2-tailed).

Fall 2005		Overall Portfolio Score	Course Grade
Overall Portfolio Score	Pearson Correlation	1	.438(**)
	Sig. (2-tailed)		.000
	N	124	123
Course Grade	Pearson Correlation	.438(**)	1
	Sig. (2-tailed)	.000	
	N	123	123

** Correlation is significant at the 0.01 level (2-tailed).

Figure 6. Correlations between overall portfolio score and course grade between fall 2004 and fall 2005.

ACKNOWLEDGMENTS

This article is part of an on-going research effort in the Department of Humanities at NJIT: Norbert Elliot taught me most of what I know and Robert Lynch supported our departmental efforts. Frank Casale, Michele Fields, Michael Kerley, Jim Lipuma, John Lyczko, Nina Pardi, and Lee Sakellarides have all contributed time and effort to this project.

REFERENCES

1. B. Huot, *(Re)Articulating Writing Assessment for Teaching and Learning*, Utah State University Press, Logan, 2002.
2. For new paradigms in writing assessment, see the following: Brian Huot, *(Re)Articulating Writing Assessment for Teaching and Learning* (Utah State

- University Press, Logan, 2002); Bob Broad, *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing* (Utah State University Press, Logan, 2003); Kathleen Blake Yancey, Postmodernism, Palimpsest, and Portfolios: Theoretical Issues in the Representation of Student Work, *College Composition and Communication*, 55:4, pp. 738-761, 2004, and Patricia Lynne, in *Coming To Terms: A Theory of Writing Assessment*, (Utah State University Press, Logan, 2004).
3. Criteria for ABET can be found on their website at: <<http://www.abet.org/forms.shtml>> August 19, 2006.
 4. Several authors write about this topic. See Bob Broad's *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing* (Utah State University Press, Logan, 2003) and Norbert Elliot's *On a Scale: A Social History of Writing Assessment in America* (Peter Lang, New York, 2005).
 5. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, *Standards for Educational and Psychological Testing*, American Psychological Association, Washington, 1999.
 6. R. Camp, New Views of Measurement and New Models for Writing Assessment, in *Assessment of Writing: Politics, Policies, and Practices*, E. M. White, D. Lutz, and S. Kamusikiri (eds.), MLA, New York, 1996.
 7. M. S. Sternglass, *Time to Know Them: A Longitudinal Study of Writing and Learning at the College Level*, Lawrence Erlbaum, Mahwah, 1997.
 8. C. M. Miller, Reuniting Wisdom and Eloquence within the Engineering Curriculum, in *Liberal Education in Twenty-First Century Engineering: Responses to ABET/EC2000 Criteria*, D. F. Ollis, K. A. Nelly, and H. C. Lugenbiehl (eds.), Lang, New York, p. 49, 2004.
 9. C. Johnson and N. Elliot, *Portfolio Assessment in Program Evaluation*, ASEE Mid-Atlantic Section 2004 Fall Conference, Engineering Workforce for the 21st Century, Digital Proceedings, October 22-23, 2004.
 10. J. Prados, G. Peterson, and L. Lattuca, Quality Assurance of Engineering Education through Accreditation: The Impact of Engineering Criteria 2000 and Its Global Influence, *Journal of Engineering Education*, 94:1, pp. 165-184, 2005.
 11. J. Coakley, *Placement in English Composition Courses: When "One-Size-Doesn't-Fit-All,"* Fourteenth International Conference on the First-Year Experience, 2001.
 12. N. Elliot, *On A Scale: A Social History of Writing Assessment in America*, Peter Lang, New York, 2005.
 13. N. Elliot, M. Kilduff, and R. Lynch, The Assessment of Technical Writing: A Case Study, *Journal of Technical Writing and Communication* 24:1, pp. 9-36, 1994.
 14. N. Coppola, Setting the Discourse Community: Tasks and Assessment for the New Technical Communication Service Course, *Technical Communication Quarterly*, 8:3, pp. 249-267, 1999.
 15. K. B. Yancey, Postmodernism, Palimpsest, and Portfolios: Theoretical Issues in the Representation of Student Work, *College Composition and Communication*, 55:4, pp. 738-761, 2004.
 16. B. Broad, *What We Really Value: Beyond Rubrics in Teaching and Assessing Writing*, Utah State University Press, Logan, 2003.
 17. N. Coppola and N. Elliot, Assessment of Graduate Programs in Technical Communication: Constructing a Model to Engage Complexity, in *Assessment in*

Professional and Technical Communication, J. Allen and M. Hundleby (eds.), SUNY Purchase, New York, forthcoming.

18. D. Dayton and S. A. Bernhardt, Results of a Survey of ATTW Members, 2003, *Technical Communication Quarterly*, 12:1, pp. 13-43, 2004.
19. H. A. Linstone and M. Turoff, *The Delphi Method: Techniques and Applications*. Available at: <<http://www.is.njit.edu/pubs/delphibook/>> August 19, 2006.

Direct reprint requests to:

Carol Siri Johnson
University Heights
New Jersey Institute of Technology
Newark, NJ 07102
e-mail: cjohnson@njit.edu