Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

ABSTRACT

WEB-BASED PORTFOLIO ASSESSMENT: AN OPEN SOURCE SOLUTION FOR PLATFORM DESIGN

by
Regina Collins

Summative assessments of student writing performance have been instrumental in the evaluation of student ability and analysis of educational programs. One method used to perform summative assessments of writing performance in post-secondary education is through the evaluation of student portfolios. Using an evidence-centered design approach, NJIT faculty researchers have developed rubrics to measure the acquired skills of students. Classroom instructors from the department meet periodically to score the students' portfolios containing constructed responses to predetermined writing tasks. The paper-based assessments are then manually key-stroked into Microsoft Excel for storage, with the scores then analyzed in SPSS and SAS.

This thesis presents the design and development of a web-based application created to enhance the portfolio assessment process and alleviate the key-stroking burden and introduction of error attendant to a paper-based portfolio scoring system. By enabling readers to rate portfolios in a communal environment in which scoring standards have been mutually established, the application ensures consistent assessment of all students in the writing program. Significantly, the application allows real-time monitoring of portfolio assessments to ensure consistency amongst readers and to immediately address portfolios requiring adjudication of discrepant scores. To ensure that the portfolio assessment platform met its full potential, both rapid prototyping and usability testing were included in the development of this application.

WEB-BASED PORTFOLIO ASSESSMENT: AN OPEN SOURCE SOLUTION FOR PLATFORM DESIGN

by
Regina Collins

A Thesis
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Professional and Technical Communication

Department of Humanities and Social Sciences

May 2010

APPROVAL PAGE

WEB-BASED PORTFOLIO ASSESSMENT:
AN OPEN SOURCE SOLUTION FOR PLATFORM DESIGN

Regina Collins

Dr. Norbert Elliot, Thesis Advisor — Date
Professor of English, NJIT

Dr. Fadi P. Deek, Committee Member — Date
Dean, College of Science and Liberal Arts,
Professor of Information Systems, Information Technology, and
Mathematical Sciences, NJIT

Dr. Andrew Klobucar, Committee Member — Date
Assistant Professor of English, NJIT

BIOGRAPHICAL SKETCH

Author: Regina Collins

Degree: Master of Science

Date: May 2010

Undergraduate and Graduate Education:

- Master of Science in Professional and Technical Communication,
  New Jersey Institute of Technology, Newark, New Jersey, 2010

- Bachelor of Science in Computer Information Systems,
  New Jersey Institute of Technology, Newark, New Jersey, 1985

Major: Professional and Technical Communication

Presentations and Publications:

Regina Collins, Nancy Steffen-Fluhr, Babajide Osatuyi, and Anatoliy Gruzd, "N is for Network: New Tools for Matting Organizational Change," Proceedings of the 2010 Women in Engineering Program Advocates Network National Conference (WEPAN 2010), Baltimore, MD, April 2010.

I have a great many people to whom I would like to dedicate this work. First and foremost are my parents. To my mother, who was ahead of her time in entering the engineering world as a woman, I thank you for being an inspiration to me; I wish that you could have lived to see this accomplishment. To my father, who always told me to follow my own path and never take no for an answer, I am forever grateful for your love and your pride in me. I wish that you, too, could be here to share the joy of my success. Aš tavę myliu. Next, I dedicate this work to my husband, my best friend, and the love of my life, Harry. Although you always say that I am smarter than you, I have never agreed with that assessment, but thank you for pushing me and telling me that the best investment we could ever make was in ourselves. Finally, I dedicate this work to my four wonderful children: Christopher, Evelyn, Emma, and Sarah. Your struggles in your own educational experiences have inspired me to persevere through my own challenges, and I hope that I have in some way inspired you as well. Learning should never end!

ACKNOWLEDGMENT

I would like to express my deepest gratitude to Dr. Norbert Elliot, my thesis advisor and mentor. In addition to being an astounding teacher, his guidance, encouragement, and faith in me have provided the scaffolding for my efforts to continually attain higher goals.

I would also like to offer my sincere appreciation to Dr. Fadi P. Deek. From his initial support of my interest in pursuing a Master's degree to his suggestions regarding this thesis, I am forever indebted to Dr. Deek for inspiring me to make my dreams reality.

I would also like to thank Dr. Andrew Klobucar, my third committee member, for his efforts throughout the development of this thesis. Dr. Klobucar's insights based on his own experiences in software development proved instrumental in the success of this thesis. As well, his patience and support relieved a great deal of stress and guilt from the process.

I am forever grateful to Dr. Nancy Coppola and to Dr. Nancy Steffen-Fluhr, both of whom supported my educational efforts both financially and, more importantly, psychologically. The confidence I gained through Dr. Coppola's nomination of me as Dr. Steffen-Fluhr's research assistant, as well as her nomination of me for the position of graduate assistant, was priceless. Working with Dr. Steffen-Fluhr introduced me not only to crucial topics relating to women in academia, but also allowed me to work with the brilliant minds of the NJIT ADVANCE grant team. Dr. Steffen-Fluhr's support and friendship have made me feel a part of the NJIT community, and I am eternally grateful to her.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The objective of this research is to design and develop a web-based, summative assessment application that automates the process of scoring student writing performance captured in portfolios. The summative assessment application follows the same evidence-centered design approach previously employed at New Jersey Institute of Technology (NJIT) using a traditional, paper and pencil method. By building upon this proven methodology of writing assessment, I can ensure that the evidence of student proficiency collected during the assessment process is valid. Furthermore, this evidence proves that the web-based portfolio assessment process accurately captures and reflects the aims of the writing program itself. As I will demonstrate, the web-based application allows summative assessments to be completed by a community of readers, thereby ensuring that all portfolio ratings benefit from the communal environment in which the raters have come to a consensus regarding scoring practices. To ensure the usability of the application, rapid prototyping of screens and subsequent usability testing were performed with both novice and experienced users.

1.1 Background Information

Students are accustomed to assessments throughout their educational experience. From grade-level report cards to standardized tests such as the SAT Reasoning Test, students, teachers, and school administrators measure success based on the results of student assessments. However, as early as 1937, concerns emerged regarding the efficacy of assessments. In an article in *School and Society*, Carl C. Brigham of Princeton University

lamented, "The pupil will gain if he is properly measured, but in the mad surge to measure two million pupils, no one is trying to describe just one pupil accurately" (p. 757). The research presented in this paper addresses the concern of Brigham by using state-of-the-art technology in the service of accurate assessment of the individual student.

1.2 Relevant Vocabulary

Before examining the effective design of assessments, it is important to understand that there are two types of assessments in use in education.

1.2.1 Formative Assessment

Formative assessment is the "assessment carried out during the instructional process for the purpose of improving teaching or learning" (Shepard, 2006, p. 627). Methods of formative assessment can include teacher observations, quizzes, and other measures that are capable of tracking the development of a student's knowledge and understanding throughout the instructional period. The information gained through formative assessment can then be used to modify instruction methods based on student feedback and progress.

1.2.2 Summative Assessment

Summative assessments, on the other hand, typically occur at the end of the instructional period and are used for the purpose of "certifying student proficiency" (Shepard, 2006, p. 627). Summative assessments are particularly important for institutions of higher learning

because successful summative assessments are a "means for creating and promoting highly effective and attractive programs for students, faculty, employers, alumni, granting agencies, and even donors" (Allen, 2004, p. 95).

1.2.3 Exit and Entrance Crunch Models of Admissions

Summative assessments enable what Elbow calls the "exit crunch" model of higher education. Instead of accepting only the most qualified applicants and weeding out undesirable students prior to admission (the "entrance crunch" model), many universities accept a broad range of students and use both formative and summative assessments to weed out those students incapable of completing the required curriculum (Elbow, 2003). Such crunch models are rejected by NJIT researchers involved in the assessment of writing ability in which summative assessments are advisory to instructors; that is, assessments performed at the end of a course are not used to prevent student matriculation through the curriculum but, rather, to allow shareholders a firm sense of programmatic student performance. At New Jersey Institute of Technology (NJIT), summative assessments provide "evidence of [our] students' learning and abilities" (Allen, 2004, p. 95); they delineate "what students learn and what they are capable of doing as a result of their educational program" (p. 96). Summative assessments allow students to display their competency in a variety of areas and then challenge their notions of their own skills by holistically examining their competencies against those of their peers.

1.3 Summative Assessment of Writing Programs

At the program level, summative assessment must measure the efficacy of an educational program in developing the particular skills required for students to succeed within that domain. However, research into existing program assessment models indicated that few, if any, specific measures of educational assessment existed. In the case of technical communication assessment at the graduate level, for instance, Coppola and Elliot (2007) developed a set of criteria from "published survey data and reports, the advice of our professional advisory board, and our own practitioner experience," resulting in the development of eight core competencies of technical communication: writing and editing, document design, rhetoric, problem solving, collaboration, interpersonal communication, specialized expertise, and technology (p. 460). Faculty work with students throughout the educational period to develop writing assignments that display mastery of the core competencies; the best of these assignments are then assembled into a portfolio for summative assessment. If the results of the assessments indicate that the general population of students is lacking in any of the competencies, faculty must reexamine the instruction of that measure within the program curriculum. A similar program has been developed at the undergraduate level by Elliot, Briller and Joshi (2007), as well as in the undergraduate technical writing curriculum (Johnson, 2006; Johnson and Elliot, 2010).

1.3.1 Portfolio Method of Summative Assessment

A critical factor in the portfolio method of summative assessment is how to accurately measure student mastery of the core competencies. Although standardized tests exist, research has shown that learning occurs in context, and therefore the most effective method for measuring learning should also occur within that same context. There is "no

perspective completely without context. There is no realm of pure exterior. All perspectives are grounded at some level in their particular" (Marcus, 1998, p. 12). Portfolio assessment – allowing students to create a portfolio of their work that they feel best exhibits their mastery of the core competencies – provides an opportunity for full construct representation; it allows experienced readers (the professors in the program) to examine students' successful acquisition of the desired target behaviors as defined by the core competencies. The use of a community of readers brought together in an environment of open discussion and mutually agreed upon standards ensures holistic assessments of not only student proficiency but programmatic success as well. This communal assessment allows "pedagogical and curricular values to be taken into account when a teaching program provides ways for faculty to interact" (Hamp-Lyons and Condon, 1993, p. 186).

1.3.2 Web-Based Portfolio Assessment

The web-based assessment application allows faculty and administrators to assess the desired target behaviors using rubrics that have already been established as valid instruments through their use in existing paper-based methods that have been used to assess both students and the writing program in general. On a broader level, the successful implementation of this assessment model in one department serves as a roadmap for implementation of similar assessment methodologies in other departments and eventually in other institutions. The web-based model is proven valid because the

data collected as part of the assessment can be used to identify "points for defensible decision making related to the curriculum, pedagogy, course sequencing, staffing, recruiting, and other matters directly related to the quality of the program" (Allen, 2004, p. 100).

CHAPTER 2

LITERATURE REVIEW

This chapter provides a brief description of the evolution of writing assessment, as well as an overview of systems currently available to enhance the assessment process.

2.1 Assessment Literature

The history of outcomes assessment may be said to begin with the formation of the Middle States Association of Colleges and Schools (MSA) in 1887; this organization would eventually create the Middle States Commission on Higher Education as one of its four components. Although initially begun by college presidents with the goal of working together to bring about legislation favorable to universities, the Association's efforts to standardize "academic credentials led to the creation of The College Board and the Carnegie Unit as ways to assure quality of academic offerings and the trustworthiness of the participating institutions" (Middle States Commission on Higher Education, 2009, p. 2) The Middle States Commission on Higher Education has similarly evolved, shifting its focus from standardization to inspection and finally to evaluation – "a qualitative assessment of achievement rather than an *a priori* commitment to a process" (Challener, 2008, p. 22).

2.2 The NJIT Writing Assessment Program

At the institutional level, one method used to achieve qualitative assessments is through portfolio assessments in writing programs. Within NJIT, there has been a coordinated effort to integrate portfolio assessment methodology into the curriculum through the inclusion of constructed response tasks as writing assignments and the creation of metrics to accurately assess student proficiency. At the undergraduate level, Johnson and Elliot (2006, 2010) developed portfolio assessment rubrics for students in the freshman level (Humanities) and junior level (Technical Communication) programs. At the graduate level, Coppola and Elliot (2007, 2010) have identified core competencies students must display, and have created rubrics to measure these competencies as evidenced in web-based student portfolios. The portfolio model of assessment in use at NJIT has further been expanded to include measurement of constructs such as information literacy (Schart, Elliot, Huey, Briller, and Joshi, 2007; Katz, Elliot, Schart, Attali, Powers, Huey, Joshi, and Briller, 2008).

Scored on a Likert-type scale, each competency measured by one of the NJIT portfolio scoring rubrics can receive a value from one to six from each reader, and each portfolio is read by at least two faculty members. (In situations where the two readers' scores are not matching or adjacent, a third reader is assigned to adjudicate.) Analysis of portfolio assessment methods revealed not only strong inter-reader reliability but also a significant relationship between student core competency scores and their overall portfolio scores. Examination of the data also identified core competencies which were not being adequately addressed by the educational program, allowing faculty and administrators to take appropriate actions to effect programmatic change.

2.3 Literature Regarding Usability

An accurate qualitative assessment cannot succeed if its deployment proves unwieldy for the users. Therefore, in designing the web-based portfolio application, close attention was paid to usability aspects of the application. In *A Practical Guide to Usability Testing*, Dumas and Redish (1999) state that usability "means that the people who use the product can do so quickly and easily to accomplish their own tasks" (p. 4). The goal of the application's usability design was to create a user interface that ensured rapid acceptance of the application as a superior alternative for performing portfolio assessments. To accomplish this goal, several usability testing methods were adapted from both traditional methods and the exploratory learning method based on the concept of pattern-based exploration (Zhao, Deek, and McHugh, in print) which encourages non-expert users to discover knowledge through usability inspection.

2.4 Outcomes Assessment in a Web-Based Environment

In her report on the state of higher education in the United States, Margaret Spellings (2006) focuses on the importance of outcomes assessment as a means of "demonstrating [higher education's] contribution to the public good" (p. 11). Outcomes assessment should be used not only to determine the "growth of student learning taking place in colleges," but also to "assess general education outcomes for undergraduates in order to improve the quality of instruction and learning" (p. 25). In short, Spellings recommends that institutions "develop interoperable outcomes-focused accountability systems designed to be accessible and useful to students, policymakers, and the public..." (p. 25).

An examination of commercial platforms designed to perform summative assessments reveals that these types of applications already exist. They include the Proficiency Profile, the Academic Profile or the Major Field Tests from the Educational Testing Service, as well as the Collegiate Assessment of Academic Proficiency from the American College of Testing. These tests are useful for demonstrating that "learning has occurred" (Middle States Commission on Higher Education, 2007, p. 30), but they are lacking in the ability to provide evidence of student mastery of particular skills. These commercial learning assessment applications can provide a general measure of student knowledge, but only evidence-centered outcomes assessments can "demonstrate that certain goals expressed in [the educational institution's] mission were achieved through exposure to the entirety of its curriculum" (Middle States Commission on Higher Education, 2007, p. 30).

Similarly, automated, web-based applications have been created specifically to test and analyze college-level writing. The iMOAT suite of web services developed at the Massachusetts Institute of Technology performs student evaluations based on essays submitted through their online system. The iMOAT system allows students to review the readings and test questions from home, take the necessary time to plan, write and edit their essay responses, and then receive detailed feedback with their results (MIT, 2003).

An examination of other web-based assessment tools reveals that a number of applications exist, but each targets a narrow aspect of learning assessment. For example, Aframe from Salmat Learning is designed specifically for corporate employee training and assessment. Vantage Learning, on the other hand, has created formative assessment tools specifically for writing programs in the K to 12 grade levels, but does not address

summative assessment. The new iCritical Thinking certification from ETS and Certiport provides a standardized method for measuring digital literacy skills. The Grady Profile, developed by Aurbach and Associates, allows teachers to create portfolios of student work and evaluate them using alternative assessment methods. Although similar in concept to the assessment methods examined in this thesis, the Grady Profile application focuses on the input of a single rater; this isolated form of assessment neglects the benefits derived by establishing a rating consensus among a group of faculty assembled to serve as a community of raters. Because the assessment instruments used in this research are evidence-centered by design, they address areas of assessment not addressed by the commercial solutions currently available. The methodology developed in this research enables evidence-centered outcomes assessment at the university level using proven instruments (rubrics) to measure students' responses to constructed response tasks. Using these rubrics, the complexities of college-level writing can receive fuller construct representation through the thorough reviews of expert readers engaged in holistic evaluations. This open forum guarantees unbiased assessments not only of each student's writing skills but also of the efficacy of the writing program itself.

Based on the evidence-centered design of the assessment and its ability to display fuller construct representation, the web-based portfolio assessment application not only provides evidence of the validity of this assessment approach to a university writing program, but also informs future research in developing evidence-centered assessment models in a web-based environment. This research creates a guide for future development of information models that will allow for the assessment of construct representations from other disciplines. Additionally, this research guides the development of formative

assessments using similar, evidence-centered models. To ensure that the application can be easily adapted by other institutions, it was developed following the guidelines of open source software. Not only does open source development reduce implementation costs, but it is widely accepted in academia. "Open source is transparent. The source code itself is viewable and available to study and comprehend. The code can be changed and then redistributed to share the changes and improvements" (Deek and McHugh, 2008, p. 1). By developing the application with a goal of providing it as an open source kernel, others will be able to build upon our work.

CHAPTER 3

MODELS OF ASSESSMENT

This chapter provides a description of the initial portfolio assessment process upon which this study was based and provides details of the benefits that occurred as a result of the new, web-based system.

The student writing tasks included in the portfolio assessments are constructed responses incorporated into the curriculum of the writing department of the university. These constructed responses are specifically designed and included in the curriculum to enable a summative assessment of students' mastery of the core competencies. In assessment, "One cannot simply construct 'good tasks' in isolation ... and hope that someone down the line will figure out 'how to score them'" (Mislevy, 2003, 2007). The rubrics in use at NJIT were specifically created by writing instructors to provide evidence of student mastery of the core competencies and have been proven as valid instruments for assessing not only student writing but the writing program itself through their use in the paper-based assessment method described in Section 2.2 above. These rubrics were integrated into the web-based solution, thereby ensuring validity of the solution and the data collected therein.

3.1 Paper-Based Assessment Model

In the previous, paper-based model of portfolio assessment in use at NJIT, a community of readers would gather in a room twice annually. With paper rubrics and pens, the faculty raters would first discuss sample portfolios to reach a scoring consensus. They would rate students' portfolios, copying the student information onto the rubric form and

then circling their selected rating responses. Every student's portfolio received two independent ratings during this assessment session.

After every student's portfolio had been rated by two raters, the rating session was ended. The faculty's rated rubrics were then given to the administrative assistants in the Department of Humanities for manual transcription of the data. Only after this manual transcription was completed were the writing assessment administrators aware of any difficulties that arose during the rating period. For example, if a student received non-matching or non-adjacent scores on any competencies measured by the rubric, the writing assessment administrator subsequently had to request a faculty member to adjudicate the discrepant scores. This adjudication was performed long after the holistic rating session had ended, meaning that the adjudicator's scores did not have the benefit of the scoring consensus established during the rating session. Similarly, if a particular rater's scores were consistently discrepant from the other raters, the writing assessment administrators were not aware of this until well after all scoring was completed. This paper-based model of portfolio assessment is shown in Figure 3.1.

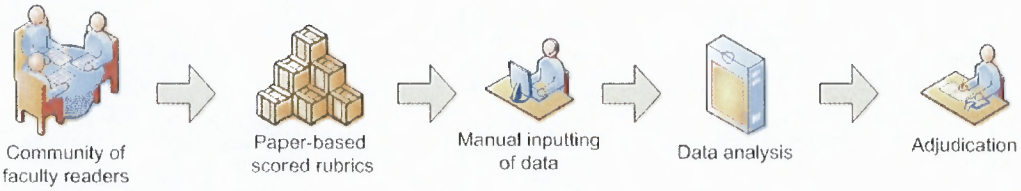

Figure 3.1 Paper-based assessment model.

3.2 Web-Based Assessment Model

Although the paper-based assessment method shown in Figure 3.1 has proven successful in capturing student proficiency and inter-reader reliability, the method has four significant disadvantages: 1) the manual transcription of the data from paper to computer is prone to human error; 2) manual data transcription is costly in terms of both time and manpower; 3) assessment administrators have no method for evaluating in real-time if there are any significant discrepancies among raters; and 4) if adjudication of a portfolio is required, this fact is not known until well after the assessment period has ended. These failures have been eliminated by the creation of the web-based application, as shown in Figure 3.2.

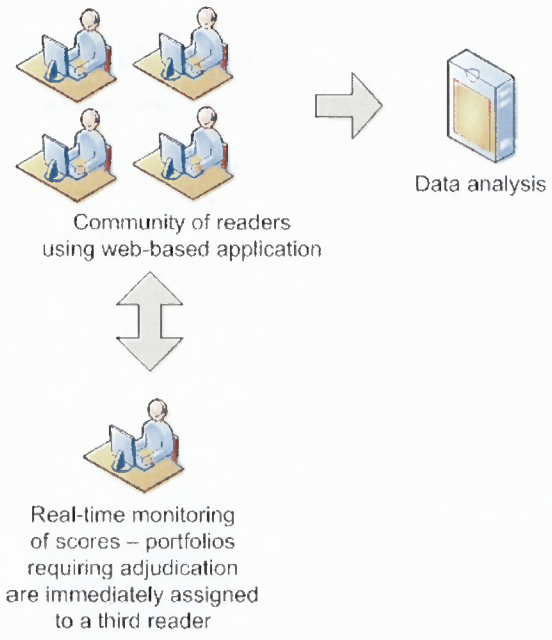

Figure 3.2 Web-based assessment model.

As with the paper-based assessment method, the web-based application was developed with a focus on evidence-centered design. In the web-based method, the readers come together at the beginning of a predetermined rating session and log into the application. All readers in the session are rating students using the same assessment rubric; in this way, prior to actual portfolio scoring, the faculty readers are able to discuss sample writing portfolios to reach a consensus regarding scoring standards. Once a consensus has been reached, the readers can select students to be rated from a drop down list. The readers then review the student's portfolio, mark their scoring selections on the web-based rubric, and commit their scores to the database. At any point, readers can modify the scores they have already assigned to students because experience has shown that the initial portfolio scores frequently require modification to maintain alignment with later scores.

During the rating session, one faculty member trained in the use of the application is designated as an administrator. The administrator can monitor the results of the rating session in real time. The application allows administrators to do the following: view reports by student (see which faculty members have rated that student and what scores have been given); view reports by rater (examine scores assigned by any particular faculty rater to see if his scores are in alignment with the established standards); assign adjudicators (see which students have discrepant scores that require adjudication and assign each student to a third reader); and identify students with incomplete ratings to ensure that all students are rated at least twice during the rating session. By enabling real-time monitoring, the application allows the administrator to ensure that all students are

fully rated prior to dismissing the faculty raters. Upon completion of the rating session, all of the scoring data (including calculations of adjudications and score totals) is complete and properly formatted for analysis.

3.3 Construct Being Measured

The construct measured in this research study was the overall, holistic experience of faculty raters using the portfolio assessment web-based application. This research examined whether the application provided an improved user experience through usability testing using a Likert scale, task-based questionnaire given to each usability tester. For each task, the construct measured the ease of completing the required task, the ease of navigating from one task to another, and the amount of information provided by supplemental text on the web pages. Through analysis of the questionnaire results, I was able to not only ascertain the overall usability of the application, but also identify areas requiring improvement.

3.4 Research Variables

The data collected as part of this research included both efficacy variables and aesthetic variables. In this particular research, efficacy was of primary importance because the application was designed to simplify a particular task – in this case, the assessment of portfolios. Efficacy variables identify the effectiveness of the application in accomplishing the assigned task, and have been separated into three major areas: task completion, navigation, and textual descriptions, as shown in Figure 3.3.

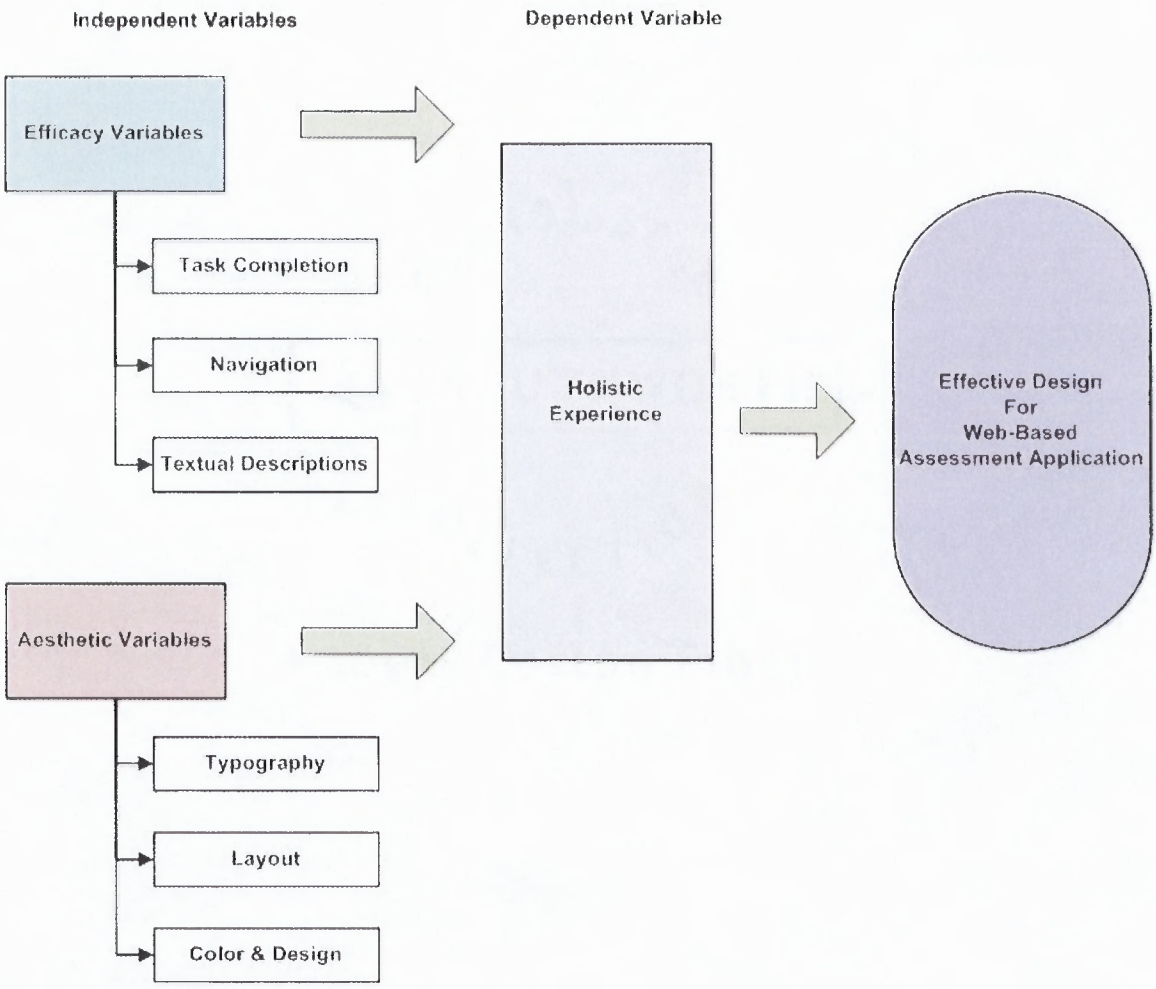

Figure 3.3 Research variable model.

Also of interest is the user's aesthetic experience in using the application. The aesthetic variables include the effective use of typography, the overall layout of the pages, and the color and design used throughout the application. However, these variables were considered secondary to the efficacy variables.

These independent variables were examined in relation to the holistic experience of using the application for portfolio assessment. A positive correlation between the independent variables and the dependent variable not only indicates the success of this research but can serve as a design model for future web-based assessment applications.

The usability design relates independent (X) variables (task completion, navigation, textual descriptions, typography, layout, color and design) to the dependent (Y) variable (the holistic score).

3.5 Validity

Every research study must establish the validity of its instruments in order for the measured data and the results to be considered accurate. The following sections describe how the content, construct, and criterion validity of this research project were established.

3.5.1 Construct Validity

The constructs used to measure the validity of this research were established based upon the extensive research literature in the field of usability and through formal consultation with experts including Susan Fowler (consultant, Fast Consulting) and Les Perelman (Director, Writing Across the Curriculum in the Program in Writing and Humanistic Studies at the Massachusetts Institute of Technology). The constructs defined the usability of the application primarily through the measurements of task completion and navigation.

3.5.2 Content Validity

The content validity of this research is based upon usability literature showing a correlation between usability measures and the overall satisfaction of users with a particular application. By establishing a positive usability experience and addressing areas identified as problematic during the application's development, we can ascertain

that usability leads to rapid acceptance and overall satisfaction with the new, web-based method of portfolio assessment.

3.5.3 Criterion Validity

Criterion validity was determined through examination of research literature regarding usability of software applications and websites. The measurement instruments developed for this research were based upon other, similar usability instruments.

3.6 Reliability

Due to the pilot nature of this research and the small faculty community of portfolio raters, testing was conducted with a small sample group of representative administrators and portfolio raters. Because the pilot testing of the application was successful, the application was put into service for the spring portfolio assessment session in May, 2010. Additional feedback from the larger-scale usage of the application will be incorporated into the application in the future.

3.7 Proving Validity and Usability

The validity of the web-based portfolio assessment application was established across several measures. First, the web-based application utilizes the same portfolio assessment rubrics already proven valid by analysis of their use in the paper-based assessment method. Further analysis of the rubrics and their competencies shows a correlation between student mastery of the identified core competencies and their overall, holistic portfolio scores (Coppola and Elliot, 2007).

By ensuring that all portfolios receive the benefit of assessment in a consensus-driven, holistic environment, and by identifying in real-time any portfolios requiring adjudication for discrepant scores, the web-based portfolio assessment application provides a highly accurate measurement of not only student proficiency but also writing program outcomes.

Usability testing throughout the development lifecycle ensured that the application was not only well-designed but that it would be readily accepted by the community of faculty readers. Usability puts a focus on the users; an application is usable when it allows users to be productive. As such, the participants in our usability studies were real users of the application (faculty members within the writing program). They were asked to perform actual tasks, and their responses were observed and recorded. During the usability study, a task-based, Likert scale survey instrument was distributed to each study participant soliciting their input on the usability of the application.

The use of open source principles enables the application to be shared with other institutions which can then adapt the instruments to their own programs. Although not developed within an open source environment, the application can be provided as a kernel in SourceForge.net so that others may use our research to develop assessments appropriate to their institutional requirements.

In general, this application serves as a proof of concept for future research into the development of a more powerful assessment platform that holds the potential to track not only student summative assessments but also formative assessments throughout the instructional phases of their studies. Such an application could provide critical information regarding a student's development of expertise in particular areas of study.

CHAPTER 4

DEVELOPMENT METHODOLOGY

The development of the web-based assessment application was performed in stages. This chapter delineates the development process and provides a detailed description of the flexibility designed into the system database to allow for future expansion of the assessment instruments in use in the portfolio assessment process.

4.1 Tools and Platforms

The development of the application was performed in a modular fashion with a focus on ensuring open source compatibility so that the application could be shared with other institutions which could adapt the instruments to their own programs. The application can be provided as a kernel in SourceForge.net so that others may use our research to develop assessments appropriate to their institutional requirements.

Development was performed locally using the XAMPP package of tools including PHP, HTML, and MySQL. After localized testing, the completed application was uploaded to the cloud using Amazon Web Services (AWS) and the Amazon Elastic Compute Cloud (EC2). By locating the application in the cloud, I eliminated the need for a dedicated server to be purchased and housed on the university campus, thereby making the web-based portfolio assessment application a cost-effective option that provides not only dependability but also flexibility. Because payment is based on usage, the cost of cloud computing is minimal. Table 4.1 defines the costs associated with cloud computing using AWS EC2.

Table 4.1 Expenses Associated with Cloud Computing

Small on-demand instance	$0.085 per hour for Linux/UNIX usage	$0.12 per hour for Windows Usage
Elastic Block Storage	$0.10 per GB-month of provisioned storage	$0.10 per 1 million I/O requests

The web-based portfolio assessment application is usage-based, meaning that it only needs to be active on the cloud during the portfolio rating sessions which occur twice per year for each of the three levels: freshman, junior, and graduate. Dedicating a full-time server to an application that will generate infrequent web traffic would create gross inefficiencies. Instead, the application resides on the cloud in Amazon's Elastic Block Storage. When an instance is required, the application administrator launches an instance for the duration of the portfolio rating session and then terminates the instance, creating a highly cost-effective solution. An added benefit of hosting this application on the cloud is the ability to expand storage and usage parameters in the future should the need arise.

4.2 Database Design

The first step in designing the web-based assessment application was to evaluate the data to be collected so that an efficient relational database could be designed. During this initial stage, data from previous assessments was reviewed to ensure proper database structure, taking into account the rules of *normalization* (Nixon, 2009). The main goal of normalization is to ensure that "each piece of information appears in the database only once" (p. 203), thereby ensuring an efficient database design.

The portfolio scoring, Likert-scale assessment rubrics were deconstructed into common elements. For example, every rubric is based upon a six-point scoring scale, and each verbal response (from Very Strongly Agree to Very Strongly Disagree) has an associated score value ranging from six to one, respectively. The table structure in the database takes advantage of these commonalities; the verbal responses and their scores are contained in a single table that can be referenced by any rubric. This modular approach simplifies the task of modifying an existing rubric or creating a new assessment rubric within the application. The database tables that define a rubric are shown in Figure 4.1.

Because of this modular structure, a portfolio assessment administrator can easily modify or create a rubric in the database. Table 4.2 explains the contents of each table and the interrelationships between the tables involved in rubric creation.

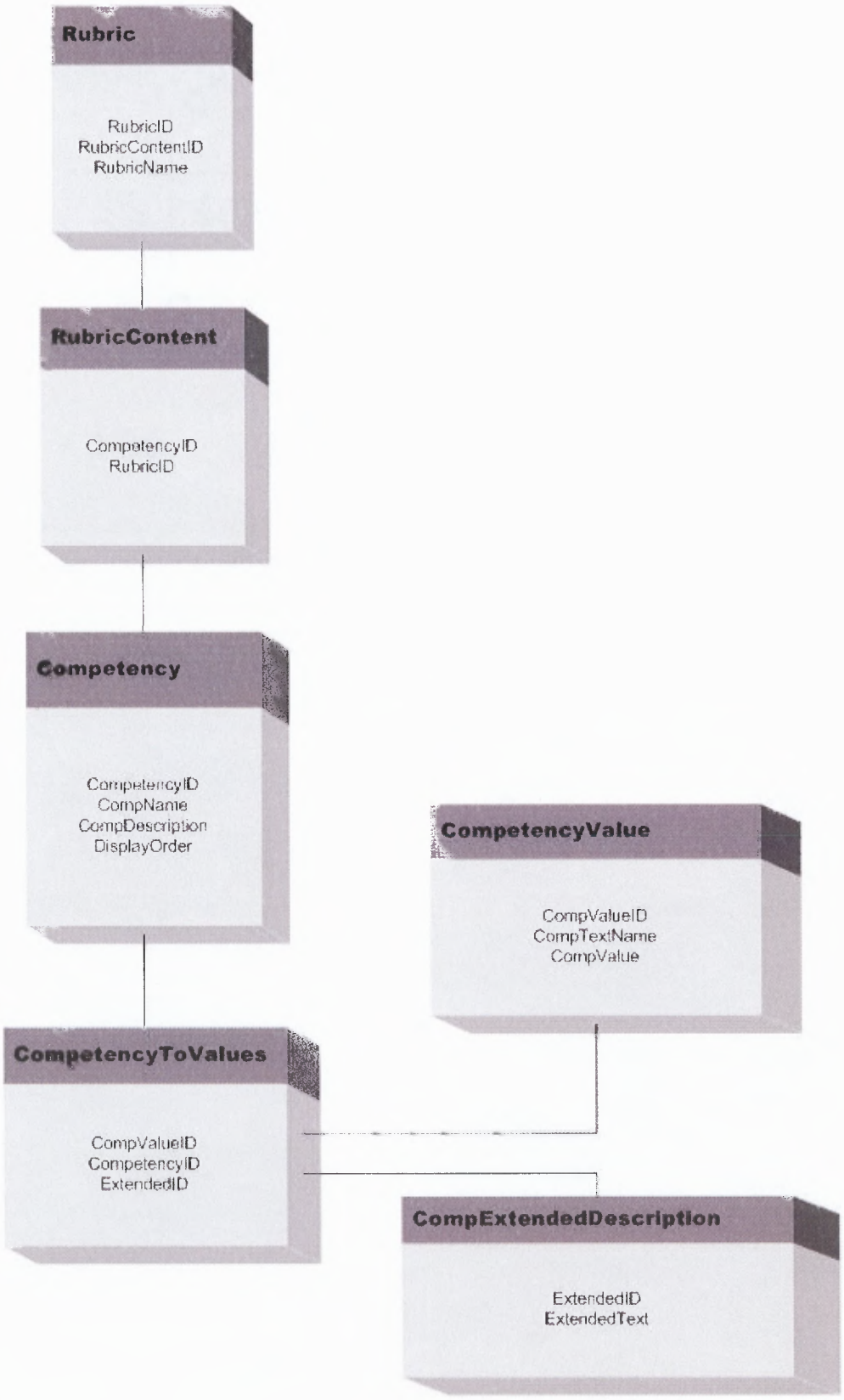

Figure 4.1 Database structure that defines each scoring rubric.

Table 4.2 Description of Database Tables for Rubric Creation

Database Table	Element	Meaning
Rubric	RubricID	Provides a unique identifier (primary key) for every rubric included in the assessment application
	RubricContentID	Provides a unique identifier (foreign key) to link data in the RubricContent table to the appropriate rubric
	RubricName	Provides a unique, user-friendly name for each rubric (e.g. HUM 101).
RubricContent	CompetencyID	Specifies which competencies are included in the content of the rubric
	RubricID	Links back to the unique rubric identifier found in the Rubric table
Competency	CompetencyID	Identifies each competency by a unique identifier (primary key). This identifier is used as a foreign key in RubricContent to associate a rubric with its competencies
	CompName	Provides a user-friendly name for each competency, such as Document Design
	CompDescription	Details the extended description provided on each rubric to explain the competency to the raters in more detail
	DisplayOrder	Stores the order in which the competency is displayed on its associated rubric

Table 4.2 Description of Database Tables for Rubric Creation (continued)

Database Table	Element	Meaning
CompetencyToValues	CompValueID	Provides a unique identifier for every competency value, from Very Strongly Agree to Very Strongly Disagree
	CompetencyID	Links the competency values to the appropriate competency
	ExtendedID	Indicates what extended text to provide for competency values. For example, all rubrics currently use extended text for the overall portfolio score rather than the Very Strongly Agree to Very Strongly Disagree values
CompetencyValue	CompValueID	Uniquely identifies the responses available for each competency (Very Strongly Agree to Very Strongly Disagree)
	CompTextName	Provides the textual names for the competency values
	CompValue	Provides the scoring value (from six to one) that corresponds with each competency value
CompExtendedDescription	ExtendedID	Indicates whether or not that competency has extended text that should be displayed
	ExtendedText	Provides the text to be displayed

4.3 Rapid Prototyping

Upon completion of the design of the database, I began the development of the actual application. Concurrently, I employed a process known as rapid prototyping (Tanik and Yeah, 1989) to ensure usability even at the earliest stages of the development cycle. By creating non-functioning HTML pages displaying the proposed user interface, I was able to solicit input from reviewers regarding the design and flow of the application prior to actual implementation.

4.4 Usability Testing

Upon completion of application development, testing with actual faculty raters was performed to ascertain the usability of the application. The application includes three assessment instruments or rubrics: Humanities 101/102, English 352, and a graduate-level rubric for the Master of Science in Professional and Technical Communication program. Each of these rubrics is included in Appendix A of this document. Usability was tested for each of these rubrics individually to ensure that any issues had been addressed.

For each assessment rubric, we invited two novice and two experienced faculty members to perform usability assessment of the application. In addition, at least one administrator was assigned to each rubric scoring session so that the administrative monitoring functions could be tested. Table 4.3 details our definition of novice and experienced users.

Table 4.3 Definitions of Novice and Experienced Usability Testers

Novice user	Faculty members who use Moodle simply as a syllabus archive or who have minimal web-based experience for instructional purposes (Note: a difficulty arose in identifying novice participants at the graduate level; the program involves distance learning, meaning that all instructors have at least some level of proficiency with Moodle and web-based learning and assessment methods.)
Experienced user	Faculty members who use Moodle as a comprehensive learning management system and, thus, have demonstrated experience using the web for instructional purposes

Each usability tester was given a task-based survey instrument based on a six-point Likert scale with which to rate the usability of the application. The application was designed in such a way that the interface for all raters, regardless of the level of the writing portfolios being assessed, is the same. Therefore, the same usability instrument was used for all rater usability tests. Similarly, the administrator interface is identical regardless of the writing portfolio level being rated; in this way, the same usability instrument was utilized for all administrative usability testers. The task-based survey instruments are included in Appendix B (Administrator Survey) and Appendix C (Rater Survey) of this document.

4.5 Data Export

One of the primary goals of the application is to monitor the data during the rating session and subsequently assemble the data for analysis immediately upon completion of the portfolio rating session. The application allows the rating session administrator to create a Results table that can be directly exported to Microsoft Excel. Prior to creating the table, the application performs all adjudication calculations to determine adjudicated scores and the final, total scores for each variable scored per student. The resulting data is in the exact format required for analysis using either SPSS or SAS. The rating session administrator can directly export the data from the application through the use of the PHPMyAdmin interface into a CSV (comma-separated values) Microsoft Excel spreadsheet.

CHAPTER 5

METHODS OF RESEARCH

The assessment of an application is a crucial step in the development process. Programmatically, rigorous testing during the development lifecycle provided evidence that the application functioned according to expectations. Throughout the application's development, sample scores were input and then analyzed to ensure adherence to the algorithms and methods used in the paper-based assessment model.

Usability testing was performed simulating a real portfolio scoring environment; novice and experienced faculty members, as well as administrators, were invited to one of three usability testing sessions. All participants were located in a single room, each at a computer workstation. Sample student portfolios were provided, and raters were asked to rate several students while the administrators monitored their progress. Situations requiring adjudication were simulated to ensure the usability of the adjudication process.

During this simulated experience, the raters were asked to score the usability of the application using a task-based survey instrument. These usability surveys are included in Appendices B and C of this document. Tables 5.1 and 5.2 provide the statement codes assigned to each survey statement, as well as the research variables measured by the statements, for purposes of data analysis.

Table 5.1 Administrator Variable Codes and Categorization by Research Variable

Statement Code	Survey Statement	Research Variables Measured
	Administrator Survey Statements	
QA1	I was able to log in to the application.	Task Completion
QA2	I understood what inputs were required for my username and password.	Textual Description, Typography
QA3	I was able to advance to the next screen.	Navigation
QA4	I understood the purpose of each menu option.	Navigation, Textual Description, Layout
QA5	There was a task I needed to perform that was not in the menu options.	Task Completion
QA6	I was able to make my selection from the menu.	Navigation
QA7	The selection did what I expected it to do.	Task Completion, Textual Description
QA8	I was able to select the appropriate rubric for the assessment session.	Task Completion
QA9	I understood the purpose of this menu option.	Textual Description

Table 5.2 Rater Variable Codes and Categorization by Research Variable

Statement Code	Survey Statement	Research Variables Measured
	Rater Survey Statements	
QR1	I was able to log in to the application.	Task Completion
QR2	I understood what inputs were required for my username and password.	Textual Description, Typography
QR3	I was able to advance to the next screen.	Navigation
QR4	I was able to enter the information for the student to be rated.	Textual Description, Layout
QR5	This page performed the task I expected.	Task Completion, Typography
QR6	I was able to advance to the next screen.	Navigation
QR7	I was able to easily select my scores on the rubric.	Task Completion
QR8	The appearance of the rubric was well organized.	Task Completion, Layout, Color and Design
QR9	The scores I selected remained marked until I chose to submit them.	Task Completion, Typography
QR10	I was able to change my selected scores as necessary prior to submission.	Task Completion, Typography
QR11	I was able to advance to the next screen.	Task Completion, Navigation

Table 5.2 Rater Variable Codes and Categorization by Research Variable (continued)

Statement Code	Survey Statement	Research Variables Measured
	Rater Survey Statements (continued)	
QR12	My task choices were clearly identified.	Navigation, Typography, Color and Design
QR13	I was able to advance to the appropriate screen.	Navigation
QR14	I understood that if I had already rated a particular student, I could modify my scores.	Task Completion
QR15	I could adjudicate students who were assigned to me	Task Completion, Textual Description, Layout

5.1 Overview of Results Analysis

The results of the usability study were analyzed based on the independent research variables identified (see Figure 3.3, Tables 5.2 and 5.3). Primary variables relating to program efficacy are task completion, navigation, and textual descriptions. Secondary variables dealing with program aesthetics are typography, layout, and color and design.

It is important to note that the number of usability testers for each usability survey (N = 4 for administrators, N = 5 for raters) was very small. Data resulting from the analysis of the usability survey results were therefore used to perform a descriptive analysis and heuristic assessment. By examining the agreement in the reported usability data, we were able to identify areas of disjunction between technology and user; this examination allowed us to pinpoint areas where usability was lacking and address those areas through modifications made to the application.

It is important to note that the usability assessment of this application is ongoing, even at the time of the present writing. Based upon feedback, I continue to revise the application to improve usability. Additional data regarding usability will be gathered when the application is deployed to perform real student portfolio assessments.

The following sections provide details regarding the analysis of the research variables based upon the data gathered through the two usability surveys: administrator and rater.

5.2 Analysis of Administrator Survey Results

In examining the inter-reader agreement of the scores provided by the administrator-level usability testers (shown in Table 5.3), it becomes apparent that administrators encountered some difficulties in using the portfolio assessment application. This is not surprising as the administrators have more tasks and menu options than the raters. Still, the disagreements between scores indicated that the administrative interface of the application was problematic and required additional usability design.

Table 5.3 Analysis of Agreement in Administrator Responses to Survey Variables

Research Variable	Matching Scores	Adjacent Scores	Outliers	Survey Statements
Efficacy				
Task Completion	7	5	4	QA1, QA5, QA7, QA8
Navigation	8	3	2	QA3, QA4, QA6
Textual Descriptions	11	1	5	QA2, QA4, QA7, QA9
Aesthetics				
Typography	3	1	1	QA2
Layout	3	0	1	QA4
Color and Design	2	1	1	QR8

For the purposes of this analysis, matching scores are usability measures that received an identical score from more than one rater. Adjacent scores differ from the matching score by one point (plus or minus), and outliers are scores that are more than one point away from the matching score. For example, raters gave usability measure QR7

the scores 6, 6, 6, 6, and 5 (N=5). This data would then be classified as four matching scores and one adjacent score with zero outliers.

Because the application will be used infrequently (twice per year), the application was designed with a goal of minimizing the need for training of portfolio assessment administrators. Instead, the platform was designed so that administrators could launch the application and follow the provided textual descriptions to understand their tasks and options within the application. An examination of the research variables shown in Table 5.3 indicates that, although the navigation and textual descriptions available to administrators appear to be adequate, there were difficulties in areas associated with task completion. This heuristic data guided subsequent revisions of the application to improve administrative usability, particularly for areas measured by the variables related to task completion.

One example of such a modification includes the method for initiating a rating session. The first implementation of the application required administrators to initiate a rating session in real time. However, data analysis revealed that there was confusion amongst the administrators in completing this process and understanding when it should be completed. Modifications were made to the application so that administrators could establish the parameters for a rating session prior to the actual session taking place. By allowing administrators to define rating sessions in advance, I also provided the opportunity for the administrators to provide a list of students to be rated, an issue that was revealed through analysis of the raters' usability survey data. Through these modifications, the raters can now log in, join the appropriate rating session, and immediately begin rating students.

Consultation with usability expert Susan Fowler elicited the following insight: "Best practices say that an infrequently used system requires more help, preferably right on the page, and the purpose of each widget needs to be very obvious and standardized." Future revisions of the application will ensure more detailed descriptions on the administrator pages of the application to guide their usage of the system. These modifications will be implemented prior to the Spring 2010 portfolio assessments, and a brief survey will be included in the application logout screen requesting additional usability feedback.

5.3 Analysis of Rater Survey Results

As previously mentioned, due to the infrequent usage of this application, an important design and usability goal was to require little to no formal documentation, particularly for the portfolio raters. The expectation is that raters who volunteer to participate in the portfolio rating sessions should be able to log into the application and immediately begin rating students. Table 5.4 highlights the inter-reader agreement for the research variables measured by the rater usability survey.

Table 5.4 Analysis of Agreement in Rater Responses to Survey Variables

Research Variable	Matching Scores	Adjacent Scores	Outliers	Survey Statements
Efficacy				
Task Completion	27	2	1	QR1, QR5, QR7, QR9, QR10, QR14
Navigation	15	7	3	QR3, QR6, QR11, QR12, QR13
Textual Descriptions	7	3	5	QR2, QR4, QR14
Aesthetics				
Typography	17	4	4	QR2, QR5, QR9, QR10, QR12
Layout	9	3	3	QR4, QR8, QR14
Color and Design	2	1	1	QR8

Areas found to be lacking in usability through analysis of the data were subsequently modified in the application. For example, usability testing revealed that the raters were concerned about the burden of entering student names and student IDs because of the opportunity for error. Based on this analysis, significant changes were made to the application's interface; instead of placing the burden of entering student information on the raters, the administrator will prepare a list of students to be rated in advance of the portfolio session. Upon joining a rating session, the rater can then see a

list of students to be rated, including the student's ID, last name, and first name, thus avoiding the opportunity for data entry errors. The application updates the drop down list throughout the rating session so that students no longer requiring rating are unavailable for selection.

In general, the agreement rates in the usability variables measuring navigation for raters indicate that the application required modification. Users experienced frustration in moving back to a previous page, or found themselves on a page from which they could not exit without completing some additional task. Based on these responses, areas of the application that revealed the highest user dissatisfaction in terms of navigation were reviewed and, where possible, alternate navigation tools were put in place. For example, if a user incorrectly input his login information, the previous version of the application required the user to manually click the back button on his browser and re-enter his information. Based on the usability survey, a capability was added so that users are automatically taken back to the login page when their login attempt fails.

The variables measured with respect to typography did not exhibit any significant correlations. From a heuristic perspective, problems were identified with aspects of typography throughout the application. For example, users frequently overlooked the instructions on the login page which described the format for their username and password. To address this concern, critical instructions were presented in a larger, darker font in the subsequent revision of the application.

The survey variables having the highest levels of agreement fall primarily into the scoring category (Survey Statements QR7 through QR11). This is important because it demonstrates that the tasks involved in scoring student portfolios were successfully

implemented. In fact, in examining the responses of the raters to the five scoring variables, it was found that the scores for all five variables were matching or adjacent and ranged from six to five on a six-point Likert scale, with two of the survey measures receiving matching scores of six from all usability testers. This is a clear indication that the raters found the application easy to use when performing the tasks involved in assessing student portfolios. Table 5.5 shows the scores for the five usability variables measuring the task of portfolio scoring.

Table 5.5 Reader Responses for Usability Variables Relating to Scoring

Reader	QR7	QR8	QR9	QR10	QR11
Reader 1	6	6	6	6	6
Reader 2	6	6	6	6	6
Reader 3	6	6	6	6	6
Reader 4	6	6	6	6	6
Reader 5	5	5	6	6	5

These scores indicate that the translation of the paper-based rubrics to the web-based application was successful. Raters were able to quickly and easily select scores for student competencies within the scoring rubric. The visual appearance of the rubrics was representative of the rubrics with which raters were already familiar, and the use of radio buttons prohibited the accidental selection of multiple scores for any single variable.

Based on these results, the scoring rubrics were not revised in subsequent versions of the application. Instead, attention was focused on areas identified as problematic by the usability survey.

CHAPTER 6

CONCLUSION

The development of a web-based portfolio assessment application provides demonstrable benefits to a variety of postsecondary shareholders: students, faculty, administrators, and accreditation agencies, as shown in Table 6.1.

Table 6.1 Shareholder Benefits

Shareholder	Benefit of Web-Based Assessment
Students	• Guaranteed consistency in portfolio reading and rating through holistic environment • Immediate availability of portfolio scores
Faculty	• More efficient and user-friendly environment for scoring portfolios • Faster feedback regarding writing program
Administrators	• Time and cost savings through elimination of transcription of data from paper-based assessments • Evidence of efficacy of faculty and curriculum
Accreditation agencies	• Proven, evidence-centered assessment that provides measurable results for the students, faculty, and curriculum at the accredited institution

This web-based portfolio assessment application bridges the gap between two distinct views of student assessment. One community believes that assessment must be performed by human readers in a personal environment (Ericcson and Haswell, 2006). The other community stresses the need for standardization through the use of machine scoring techniques (Shermis and Burstein, 2003). The web-based portfolio application achieves the goals of both communities: it enhances and simplifies the human scoring process through automation of the cumbersome tasks of input, transcription and analysis while still maintaining the human element.

By allowing human readers to come together in a holistic scoring environment, the web-based portfolio application provides the following benefits: it enables standardization of scores through a rating consensus reached by the faculty raters; it uses technology to simplify the raters' scoring tasks by providing point-and-click scoring rubrics; it allows administrators to monitor the progress of the scoring in real-time so that rating inconsistencies can be immediately identified and addressed; and it allows portfolios requiring adjudication due to discrepant scores to be immediately identified and assigned to a third reader within the same holistic scoring environment. The web-based portfolio application enhances and simplifies, but does not reduce, the construct crucial to student writing assessment.

The web-based portfolio application was developed using commonly accessible software (PHP, MySQL, and HTML) utilizing open source principles so that the source code could be provided as a kernel on SourceForge.net. By sharing this research through

the open source community, other institutions of higher education can expand upon this work to enable web-based assessments of their own writing programs or other programs within their institutions

By performing usability analysis with a small group of respondents, this research defines a process through which data can be analyzed not to provide proof of correlation but to perform heuristic analysis of an application's usability. Through a careful examination of the usability data, we were able to identify problematic areas in the application and address the usability issues through rapid deployment of application modifications. In situations where the immediate group of shareholders is relatively small, this type of small group usability testing and subsequent inferential analysis can prove highly effective in directing software development efforts towards the areas in most need of modification, thereby ensuring the ultimate usability and success of the application.

APPENDIX A

PORTFOLIO ASSESSMENT RUBRICS

The current implementation of the web-based portfolio assessment application utilizes the same rubrics that have been developed, tested, and put into use in the paper-based portfolio assessment process. These rubrics were replicated in the web-based application to ensure validity and reliability with the paper-based system through the use of these proven instruments for student writing assessment. This appendix includes each of these rubrics: Humanities 101, 102; English 352 (Technical Communication); and Master of Science in Professional and Technical Communication (MSPTC).

A.1 Humanities 101-102 Assessment Rubric

HUM 101-102 Writing, Speaking Thinking *Portfolio Assessment*

Student's Name: _____ Student ID:_____

Reader's Name: _____ Course and Section:_____

Provide an analytic reading in which you focus on the FOUR traits identified below:
1. Critical Thinking 2. Revising and Editing 3. Written Language 4. Information Literacy

1. **Critical Thinking** The contents of the portfolio demonstrate that the student has thought critically in preparing written assignments.

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

2. **Revising and Editing** The contents of the portfolio demonstrate that the student has drafted and successfully revised papers before they were submitted.

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

3. **Written Language**

A) Content and Organization: The contents of the portfolio demonstrate that the student writes with purposeful organization and makes connections between ideas that progress clearly from beginning to end.

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

B) Sentence Construction and Mechanics: The contents of the portfolio demonstrate that the student writes clear, well-formed sentences, using accurate grammar, punctuation and spelling.

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

Holistic Portfolio Score Provide an overall, holistic impression of the portfolio.

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree
The materials in the portfolio demonstrate *excellent* work in the class.	The materials in the portfolio demonstrate *very good* work in the class.	The materials in the portfolio demonstrate and an *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 near failure.*	The materials in the portfolio demonstrate work that is at a level of *failure.*

A.2 English 352 Assessment Rubric

English 352 (Technical Communication) Portfolio Assessment – Spring 2009

Student's Name _____

Reader's Name _____

Web Page

1. The web page is clear and navigable.

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

Writing and Editing

2. The contents of this portfolio exhibit clear style (readable, concise, cohesive).

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

3. The contents of this portfolio demonstrate accurate language usage (grammar, punctuation, spelling).

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

Substance and Content

4. The contents of this portfolio exhibit clear understanding of assignments.

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

5. The contents of this portfolio demonstrate accurate, thorough, relevant, and coherent content and ideas.

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

Audience Awareness

6. The contents of this portfolio demonstrate that the author can adapt tone for audience.

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

Document Design

7. The contents of this portfolio demonstrate cohesion by graphic means (headings, white space) in documents.

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

Information Literacy

8. Citation: This portfolio includes sources that are documented so that the original source can easily be found.

Very Strongly Agree	Strongly Agree	Agree	Disagree	Strongly Disagree	Very Strongly Disagree

Overall Portfolio Score:

9. The materials in this portfolio demonstrate class work that is:

Superior	Very Good	Average	Below Average	Near Failure	At Failure

A.3 Master of Science in Professional and Technical Communication Assessment Rubric

MSPTC Online Portfolio Assessment Spring 09

Reader's Name_____ Student's Name _____

Writing and Editing

The contents of the portfolio demonstrate that the student has competent writing and editing skills, as described in the assessment matrix.

Very Strongly Agree❺	Strongly Agree❹	Agree❸	Disagree❷	Strongly Disagree❷	Very Strongly Disagree ❶

Document Design

The contents of the portfolio demonstrate that the student has competent document design skills, as described in the assessment matrix.

Very Strongly Agree❺	Strongly Agree❹	Agree❸	Disagree❷	Strongly Disagree❷	Very Strongly Disagree ❶

Rhetoric

The contents of the portfolio demonstrate that the student has competent rhetorical skills, as described in the assessment matrix.

Very Strongly Agree❺	Strongly Agree❹	Agree❸	Disagree❷	Strongly Disagree❷	Very Strongly Disagree ❶

Personal Traits, Work Skills, Problem Solving

The contents of the portfolio demonstrate that the student has competent work and problem solving skills, as described in the assessment matrix.

Very Strongly Agree❺	Strongly Agree❹	Agree❸	Disagree❷	Strongly Disagree❷	Very Strongly Disagree ❶

Collaboration and Team Work

The contents of the portfolio demonstrate that the student has had experience working in teams.

Very Strongly Agree❺	Strongly Agree❹	Agree❸	Disagree❷	Strongly Disagree❷	Very Strongly Disagree ❶

Oral or Interpersonal Communication

The contents of the portfolio demonstrate that the student has competent oral or interpersonal communication skills, as described in the assessment matrix.

Very Strongly Agree❺	Strongly Agree❹	Agree❸	Disagree❷	Strongly Disagree❷	Very Strongly Disagree ❶

Specialized Expertise

The contents of the portfolio demonstrate that the student has competent research skills.

Very Strongly Agree❺	Strongly Agree❹	Agree❸	Disagree❷	Strongly Disagree❷	Very Strongly Disagree ❶

Technology

The contents of the portfolio demonstrate that the student has proficiency with technology.

Very Strongly Agree❺	Strongly Agree❹	Agree❸	Disagree❷	Strongly Disagree❷	Very Strongly Disagree ❶

Overall Portfolio Score

The materials in the portfolio demonstrate superior work in the program. ❺	The materials in the portfolio demonstrate very good work in the program.❹	The materials in the portfolio demonstrate average work in the program❸	The materials in the portfolio demonstrate below average work in the program.❷	The materials in the portfolio demonstrate work that is at a level of near failure in the program.❷	The materials in the portfolio demonstrate work that is at a level of failure in the program.❶

APPENDIX B

USABILITY SURVEY FOR ADMINISTRATORS

Usability of the application was tested with novice and experienced faculty members using a task-based survey instrument. Because the tasks for administrators are different than those for raters, two separate survey instruments were created. This appendix documents the task-based survey distributed to administrators.

B.1 Protocol for Usability Testing

An integral part of application development is ensuring usability. The following task-based protocol was used to examine the usability of the web-based portfolio assessment application. The protocol begins with a pretest questionnaire to collect data about the participants of the usability study. The subsequent questions deal with particular tasks that had to be performed during the testing.

The protocol contains two sections: a pre-test questionnaire and a task-based usability questionnaire. Your input is greatly appreciated.

B.2 Pretest Questionnaire

1. What is your name: _____

2. What is your gender: □ Male □ Female

3. What is your current age: _____

4. Are you of Hispanic, Latino, or Spanish origin?

   □ No, not of Hispanic, Latino, or Spanish origin

   □ Yes, Mexican, Mexican Am., Chicano

   □ Yes, Puerto Rican

   □ Yes, Cuban

   □ Yes, another Hispanic, Latino, or Spanish origin:

   (please specify: _____)

5. What is your race?

   □ White

   □ Black, African Am., or Negro

   □ American Indian or Alaska Native

   □ Asian Indian

   □ Chinese

   □ Filipino

   □ Other Asian

   (please specify: _____)

   □ Japanese

   □ Korean

   □ Vietnamese

□ Native Hawaiian

□ Guamanian or Chamorro

□ Samoan

□ Other Pacific Islander

(please specify: _____)

□ Some other race

(please specify: _____)

6. What is your job title: _____

7. How long have you been doing this work: _____

8. Please rate your skills in using software such as the Moodle Course Management System:

□ Very experienced (use most or all of the functionality available)

□ Moderately experienced (use some of the functionality available)

□ Experienced (use the basic functionality available)

□ Novice (have not used or used only in a limited capacity)

9. Have you participated in portfolio assessments prior to this session?

□ Yes □ No

If yes, how frequently? _____

B.3 Administrator Usability Protocol

Login:

QA1: I was able to log in to the application.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA2: I understood what inputs were required for my username and password.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA3: I was able to advance to the next screen.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

Administrator Functions:

QA4: I understood the purpose of each menu option.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA5: There was a task that I needed to perform that was not in the menu options.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA6: I was able to make my selection from the menu.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA7: The selection did what I expected it to do.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

Set Assessment Parameters:

QA8: I was able to select the appropriate rubric for the assessment session.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA9: I understood the purpose of this menu option.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

Assign Adjudicator:

QA10: I understood what an adjudicator was.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA11: The supporting text provided an accurate description of the adjudication parameters.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA12: I did not want to see students already adjudicated in the list.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA13: I was able to select a student for adjudication.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

View Reports:

QA14: The report descriptions gave me an understanding of what each report would display.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA15: I was able to select my desired report type.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA16: I would prefer to have more detailed descriptions of each type of report on this page.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA17: I was able to advance to the next screen.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

View Scores by Rater:

QA18: I was able to select a rater for analysis.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA19: I would prefer to see a list of raters than to manually enter the rater information.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA20: This report did not provide the information I expected it would.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA21: I was able to advance to the next screen.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

View Scores by Student:

QA22: I was able to select a student for analysis.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA23: I would prefer to see a list of students than to manually enter the rater information.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA24: This report did not provide the information I expected it would.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA25: I was able to advance to the next screen.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

View Scores by Variable:

QA26: I was able to select a variable for analysis.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA27: The information provided by the supporting text was sufficient to help me understand the report.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA28: This report did not provide the information I expected it would.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA29: I was able to advance to the next screen.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

View Records Requiring Adjudication:

QA30: I was able to understand the information presented by this report.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA31: This report did not provide the information I expected it would.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA32: I did not understand the option to sort by rater names.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA33: I was able to advance to the next screen.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

Export Data for Analysis:

QA34: I was able to export the data to a specified file name.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QA35: I was able to advance to the next screen.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved:

_____

APPENDIX C

USABILITY SURVEY FOR RATERS

Usability of the application was tested with novice and experienced faculty members using a task-based survey instrument. Because the tasks for administrators are different than those for raters, two separate survey instruments were created. This appendix documents the task-based survey distributed to raters.

C.1 Protocol for Usability Testing

An integral part of application development is ensuring usability. The following task-based protocol was used to examine the usability of the web-based portfolio assessment application. The protocol begins with a pretest questionnaire to collect data about the participants of the usability study. The subsequent questions deal with particular tasks that had to be performed during the testing.

The protocol contains two sections: a pre-test questionnaire and a task-based usability questionnaire. Your input is greatly appreciated.

C.2 Pretest Questionnaire

1. What is your name: _____
2. What is your gender: □ Male □ Female
3. What is your current age: _____
4. Are you of Hispanic, Latino, or Spanish origin?

   □ No, not of Hispanic, Latino, or Spanish origin

   □ Yes, Mexican, Mexican Am., Chicano

   □ Yes, Puerto Rican

   □ Yes, Cuban

   □ Yes, another Hispanic, Latino, or Spanish origin:

   (please specify: _____)
5. What is your race?

   □ White

   □ Black, African Am., or Negro

   □ American Indian or Alaska Native

   □ Asian Indian

   □ Chinese

   □ Filipino

   □ Other Asian

   (please specify: _____)

   □ Japanese

   □ Korean

   □ Vietnamese

□ Native Hawaiian

□ Guamanian or Chamorro

□ Samoan

□ Other Pacific Islander

(please specify: _____)

□ Some other race

(please specify: _____)

6. What is your job title: _____

7. How long have you been doing this work: _____

8. Please rate your skills in using software such as the Moodle Course Management System:

□ Very experienced (use most or all of the functionality available)

□ Moderately experienced (use some of the functionality available)

□ Experienced (use the basic functionality available)

□ Novice (have not used or used only in a limited capacity)

9. Have you participated in portfolio assessments prior to this session?

□ Yes □ No

If yes, how frequently? _____

C.3 Rater Usability Protocol

Login:

QR1: I was able to log in to the application.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QR2: I understood what inputs were required for my username and password.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QR3: I was able to advance to the next screen.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

Input student to rate:

QR4: I was able to enter the information for the student to be rated.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QR5: This page performed the task I expected.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QR6: I was able to advance to the next screen.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

Student scoring:

QR7: I was able to easily select my scores on the rubric.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QR8: The appearance of the rubric was well organized.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QR9: The scores I selected remained marked until I chose to submit them.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QR10: I was able to change my selected scores as necessary prior to submitting them.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QR11: I was able to advance to the next screen.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

Select next action:

QR12: My task choices were clearly identified.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QR13: I was able to advance to the appropriate screen.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QR14: I understood that if I had already rated a particular student, I could modify my scores.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

QR15: I could adjudicate students who were assigned to me.

6	5	4	3	2	1
Very strongly agree	Strongly Agree	Agree	Disagree	Strongly disagree	Very Strongly Disagree

If you answered 3, 2, or 1 to any of the above questions, please provide suggestions on how the application could be improved: _____

WORKS CITED AND CONSULTED

Allen, Jo. (2004). The Impact of Student Learning Outcomes Assessment on Technical and Professional Communication Programs. *Technical Communication Quarterly 13.1,* 93-108.

Brennan, R. L. (2006). Perspectives on the evolution and future of educational measurement. In R. L. Brennan (Ed.). *Educational measurement* (4th ed., pp. 1-16). Westport, CT: American Council on Education and Praeger.

Coppola, Nancy W., and Norbert Elliot. (2007). A Technology Transfer Model for Program Assessment in Technical Communication. *Technical Communication 54.4*, 459-74.

---. Assessment of Graduate Programs in Technical Communication. *Assessment in Technical and Professional Communication.* M. Hundleby & J. Allen (Eds.) Amityville, NY: Baywood. 127-160.

Deek, Fadi P., and Norbert Elliot. (in press). Validation in Post-Secondary Assessment: Toward a Unified Theory. *Proceedings, Center for Program and Learning Assessment, Program and Learning Assessment Conference*. November 6, 2009.

Deek, Fadi P., and James A. M. McHugh. (2008) *Open Source: Technology and Policy*. Cambridge: Cambridge UP.

Dumas, Joseph S. and Janice C. Redish. (1999). *A Practical Guide to Usability Testing* Revised Ed. Portland: Intellect Books.

Elbow, Peter. (2003). Directed Self-Placement In Relation To Assessment: Shifting the Crunch from Entrance to Exit. *Directed Self-Placement: Principles and Practices.* Ed. Daniel J. Royer and Roger Gilles. Cresskill: Hampton Press, Inc. 15-30.

Elliot, N., V. Briller and K. Joshi. (2007). Analytic Portfolio Assessment: A Program Development Model. *Journal of Writing Assessment 3,* 5-30.

Ericcson, Patricia, and Richard Haswell. (2006) *Machine Scoring of Student Essays: Truth and Consequences*. Utah: Utah State UP.

Fowler, Susan. "Re: Usability testing of web-based portfolio assessment application." Message to the author. 26 April 2010. E-mail.

Hamp-Lyons, Liz, and William Condon. (1993). Questioning Assumptions about Portfolio-Based Assessment. *College Composition and Communication 44.2,* 176-190.

Johnson, C. (2006). A Decade of Research: Addressing Change in the Technical Communication Classroom Using Online Portfolios. *Journal of Technical Writing and Communication 26.4*, 413-31.

Johnson, C., and N. Elliot. (2010). Undergraduate Technical Writing Assessment: A Model. *Programmatic Perspectives.*

Katz, I. R., N. Elliot, D. Scharf, Y. Attali, D. Powers, H. Huey, K. Joshi, and V. Briller. (2008). *Information Literacy Assessment: Case Study of a Multi-Method Approach* (ETS Research Report RR-08-33). Princeton, NJ: Educational Testing Service.

Landauer, Thomas K., Peter W. Foltz, and Darrell Laham. (1998). An Introduction to Latent Semantic Analysis. *Discourse Processes 25,* 259-84.

Marcus, K. A. (1998). Science, measurement, and validity: Is completion of Samuel Messick's synthesis possible? *Social Indicators Research, 45*, 7-34.

Massachusetts Institute of Technology. (2003). iCampus Project: IMoat: The iCampus/MIT Online Assessment Tool. web.mit.edu/imoat.

Middle States Commission on Higher Education. (2007). *Student Learning Assessment: Options and Resources* 2nd Ed. Philadelphia: Middle States Commission on Higher Education.

Nixon, Robin. (2009). *Learning PHP, MySQL & JavaScript*. Sebastopol: O'Reilly Media, Inc.

Royer, D.J., & Gilles, R. (1998) Directed self-placement: An attitude of orientation. *College Composition and Communication, 50*, 54-70.

Scharf, D., N. Elliot, H. Huey, V. Briller, and K. Joshi (2007). Direct Assessment of Information Literacy Using Writing Portfolios. *The Journal of Academic Librarianship, 22*, 462-78.

Shermis, and Burstein. (2003). *Automated Essay Scoring: A Cross-Disciplinary Perspective*. Mahwah, NJ: Lawrence Erlbaum.

Spellings, Margaret. (2006). *A Test of Leadership: Charting the Future of U.S. Higher Education.*

Tanik, Murat M., and Raymond T. Yeh. (1989) Rapid Prototyping in Software Development. *Computer 22.5*, 9-11.

Zhao, Luyin, Fadi P. Deek, and James A. McHugh. (in print). Exploratory Inspection – A User-Based Learning Method for Improving Open Source Software Usability. *Journal of Software Maintenance and Evolution: Research and Practice.*