A STANDARDS-BASED GRADING AND REPORTING TOOL FOR FACULTY: DESIGN AND IMPLICATIONS

By

ALAA M. SADIK

Assistant Professor, Instructional Technology at Sultan Qaboos University, Oman.

ABSTRACT

The use of standard-based assessment, grading and reporting tools is essential to ensure that assessment meets acceptable levels of quality and standardization. This study reports the design, development and evaluation of a standards-based assessment tool for the instructors at Sultan Qaboos University, Sultanate of Oman. The Rapid Applications Development Model was implemented to develop early versions of the assessment tool, called RealGrade. The Grading tool Usability Questionnaire and a series of individual interviews were used to measure participants' reactions toward the usability of RealGrade and determine the extent to which the prototype is usable. The results revealed that participants found the RealGrade effective and efficient in facilitating the process of standards-based assessment and communicating grades with students at the University. In addition, they favored the design, flexibility and ease of use of RealGrade. Further examinations of mean differences among participants according to their computer experience and teaching experience were conducted.

Keywords: Standards-based Assessment, Assessment Tool, Usability Evaluation.

INTRODUCTION

Assessment is defined as the systematic collection of information about student performance in order to inform decisions about how to improve learning (Walvoord, 2004). Grades are the standard method for reporting student performance across universities. They represent the essence that is education and serve as a mechanism for communication between instructors and students (Hills, 1991). To offset the ambiguity of information communicated via grades, instructors usually follow a process to determine the nature and number of assessments on which to base grades, select the weight to give each assessment, and set the performance standard for each grade (Oosterhof, 1994). Research has identified four major functions of grades: administrative, guidance, information, motivation and discipline (MacCormack, 2001 and Scacchi, 2000). However, the variance in assessment methods and grading practices makes assessment information used for all four of these major functions suspect. According to Ebel and Frisbie (1986), grades obtain their meaning from one or more of the following three measurement sources: (i) a comparison of a student's achievement with some absolute or relative standard; (ii) the quality of performance with respect to either amount of effort or achievement; or (iii) the amount of knowledge or learning attributable to the course. Of these three sources, research findings support the first method of comparing achievement to some standard (Khattri, Kane, & Reeve, 1995; Burger, 1998; Guskey, 2001).

Currently, there are an increasing number of reasons for universities to teach using a standards-based approach. Khattri, Kane, and Reeve (1995) indicated that performance assessments have a positive influence on education and provide developmentally appropriate frameworks for evaluation(*). Performance standards indicate what is required to meet content standards as well as the quality of achievement that is deemed acceptable (Burger, 1998). Therefore, instructors need to attach content and performance standards to assignments and activities to observe trends regarding how students perform over time for each standard; this will help instructors accurately assess student proficiencies within each standard. In addition, instructors need a comprehensive grading and

(*)The term "standard" is used synonymously to refer to curriculum standards, content standards, and performance standards. Curriculum standards describe what should take place in the classroom; as such, they address instructional techniques and recommended activities and various modes of presentation. Content standards describe what students should know or be able to do.

reporting system that shows how students are measuring up to standards (Gus key, 2001).

At the same time, instructors and students like reporting formats that are easy to understand. They do not want reports that are difficult to read and analyze (Burger, 1998). Although advances in computerized reporting forms allow instructors to provide such simple, individualized reports, few instructors have taken up the challenge. With the numerous advances in computer software, instructors' utilities, such as standards-based grading tools, can yield information about the strengths and weaknesses of students in particular content and skill areas as well as ensure that this information is provided to students in a useful and comprehensible manner (Gus key, 2001). In other words, there is a need for a system that shows what students know in relation to course standards rather than the current system where grades do not always relate to course content.

Assessment tools, whether traditional or electronic, are the official documents for recording student grades and are a primary source of student grade data. Usually, electronic assessment tools provide information about the total number of student scores used to aggregate each student's grades, the activities graded, the system used to record scores, and a summative grade for each student (Reed, 1996). In addition, these tools may provide relief to instructors who find themselves entrenched by tracking student performance, recording results of academic activities, calculating grades and reporting exam results (Roblyer, Edwards, & Havrileck, 1999).

Recent research in human-computer interaction emphasizes the importance of usability as a major element in software design and as a strong indicator of the overall acceptability of software (Preece, Rogers, and Sharp, 2002; Rozanski and Haake, 2003). Traditionally, software usability has been defined as a quality attribute that assesses how easy software is to use (Nielsen, 2003). The ISO 9241 guide on usability provided the most accepted and adopted definition in the literature. According to ISO 9241 (1998), usability is defined as the extent to which a system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. Effectiveness is defined as the accuracy and completeness with which specified users can achieve specified goals in particular environments. Indicators of effectiveness include quality of solution and error rates. Efficiency is the resources expended in relation to the accuracy and completeness of goals achieved. Indicators of efficiency are the completion time and ease to learn. Satisfaction is the users' comfort with and positive attitudes towards the use of the system. Users' satisfaction can be measured by attitude rating scales.

In usability evaluation, attention is given to ensuring not only that software works as intended but also that the user-interface is effective so the user can concentrate on the process instead of the interface (Bevan, 2001). In addition, attention should be paid to user satisfaction as a particular aspect of usability. Rubin (1994) highlights some aspects that can be used to measure user satisfaction, such as perceived usefulness and how well software matches expectations.

However, studies directly addressing the development and evaluation of standards-based grading software at the university level are virtually non-existent. In addition, although many commercial assessment tools are available, none have been developed for specific instructors' needs and differences in mind. The majority of these applications are designed for either school teachers or instructors at a specific university system. Therefore, instructors are being challenged by administrative demands of processing standards-based assessment and are not able to integrate any of these applications into their grading practices. This situation has placed an emphasis on the need to develop and evaluate a software for Instructors at Sultan Qaboos University (SQU) that could result in a usable standards-based assessment and grading tool.

Problem of the study

SQU is a cross-cultural organization and the largest academic community in Oman that consists of nine colleges and brings together hundreds of instructors from around the world. More than 2500 students were enrolled in the 2009-2010 academic year. The University is committed

to improving the quality and understanding of the education and social services provision in Oman, and it is involved in a number of initiatives and programs working to advance teaching practices. However, a wide variance was observed across the SQU with regard to the calculation of student grades. There is a wide variety of ways in which instructors score, tabulate grades and prepare report cards. In addition, many instructors believe that preparing, scoring, grading, and reporting student academic performance are each extremely exhausting and difficult tasks, requiring the tabulations of an entire term to be done traditionally, even with the aid of spreadsheets. Many instructors have found that it is difficult to perform daily tracking and standards-based assessment according to the university quality assurance policy and requirements.

Research questions

This investigation aims to increase the understanding of the usability of the standards-based assessment tool used by instructors and seeks to answer the following questions:

- How effective (useful) is the assessment tool in documenting, grading and reporting student performance?
- How efficient (easy to use) is the assessment tool as perceived by users?
- What is the overall satisfaction of users toward the use of the assessment tool?
- What individual difference variables influence instructors' perceived usability of the assessment tool?

Purpose of the Study

The main purpose of this study was to design, develop and evaluate a standards-based assessment tool for instructors at SQU. This tool should assist instructors in documenting, managing and communicating student performance based on content and performance standards. In addition, the grading tool should accommodate the cultural and technical differences among instructors, as well as the requirements of standards-based assessment at the University.

Significance of the Study

Because of issues such as differences in the traditional and electronic methods used in evaluating student performance, developing and evaluating a standards-based assessment tool should bring consistency to these practices. Not only does using an electronic assessment tool promote consistency, but it also assists in promoting professionalism in the documentation process throughout SQU. SQU is putting forth great effort to integrate technology in various ways in this electronic and digital era as it moves toward accreditation. The use of a standard assessment, grading and reporting tool is essential to ensure that assessment meets acceptable levels of the digital age, quality and standardization, which are basic requirements for accreditation.

Method

Development of the assessment tool

The rapid applications development model (RAD) was found the most efficient model of software development relative to other models. It offers a framework within which quality software can be developed on time and within budget, particularly for educational institutions (Rushby, 1997). The RAD model allowed the developer to rapidly construct the primitive version of software system that users can evaluate. User evaluations can then be incorporated as feedback to refine the emerging system specifications and designs (Scacchi, 2001).

Based on the RAD model, a user-needs analysis for the assessment tool was carried out first. The main purposes of the analysis were to ensure faculty involvement throughout the development process, determine the gap between the existing grading skills and knowledge of faculty and those that are needed for the assessment tool, and define the grading requirements that the assessment tool must fulfill. Consequently, a series of individual interviews and focus groups were conducted with faculty across the University to investigate these issues. Example questions included the following:

- What are the problems you face in documenting, manipulating and reporting students' standards-based grades manually, using spreadsheets, or other types of assessment tools?
- What are the functions and features you expect in a standards-based assessment tool for grading and reporting student performance at SQU?

Interviews and focus groups revealed that the task of grading and reporting student performance is very time consuming even with the use of Excel spreadsheets. They indicated that although Excel is a powerful application, it is a very frustrating grading tool especially for faculty trying to combine and weight course activities. An Excel user commented that "if you make an error in your grading formula, every single calculation done on that spreadsheet will be wrong". In addition, many faculty members indicated difficulty importing and editing class lists from the University Student Information System (SIS) to generate attendance sheets and report grades. Information from the SIS is usually in CSV and HTML formats. Overall, instructors believe that the proposed assessment tool should be able to do the following:

- Support the University grading scale and generate students' grades automatically;
- Enable qualitative assessment of student performance;
- Attach content and performance standards to grade sheets;
- Be compatible with the University Student Information System, from importing class lists to submitting final grades;
- Track student attendance with absence warning indicators;
- Have a built-in e-mail function for communicating grades with students;
- Facilitate total point and percent weight of scores and assignments;
- Generate course statistics for distributions, correlations and variances;
- Provide attractive print-outs for grade sheets, attendance sheets, course statistics, and grade reports.

To determine whether the prototype met the needs and expectations of faculty and to collect user-performance and satisfaction data at an early stage in the grading tool development, a series of tryouts were conducted using one-to-one and small groups of target users (5-10 users). Participants were selected from University instructors who volunteered; the computer experience of the volunteers varied. Each tryout was carried out for one week. Observations showed that many participants were confused even by basic operations in the grading tool (e.g., importing class lists). They had difficulty understanding what the product could do for them, where to go to perform an operation, and how to perform that operation once they found it. Many users suggested that simplifying the grading tool would be a useful way to satisfy and attract new users.

The interviews highlighted many specific issues related to user-interface design, data-inputs and outputs, dealing with student information and data, scoring and grading academic activities, importing and exporting files, and weighting scores. The prototype was modified and improved in light of the above feedback, and more individual and group tryouts were carried out to make sure that the assessment tool performed the planned functionality in the best way possible. Various issues highlighted by participants were considered as valuable feedback used to improve the prototype, called RealGrade. Figure 1 shows RealGrade main spreadsheet-like users interface and statistics window.

Finally, various importing and exporting functions were provided to assist users in importing student information directly from the university Student Information System (SIS), Excel spreadsheets, or a course management system (Moodle) and in exporting grades directly to SIS or Excel. The analytical functions implemented in RealGrade included basic statistical analyses (distribution of grades, mean, minimum, maximum, standard deviation, and variance) and graphical analyses (histogram, stacked line, Skewness, and kurtosis). RealGrade also provided many functions to export, print, upload and communicate student performance and grades.

Sample

Preece, Rogers and Sharp (2002) indicated that in usability evaluation, participants must be appropriate users who represent the target user population. Therefore, an email message was sent to a random sample of instructors at different colleges across SQU (N=340) asking them to participate in a study investigating the usability of

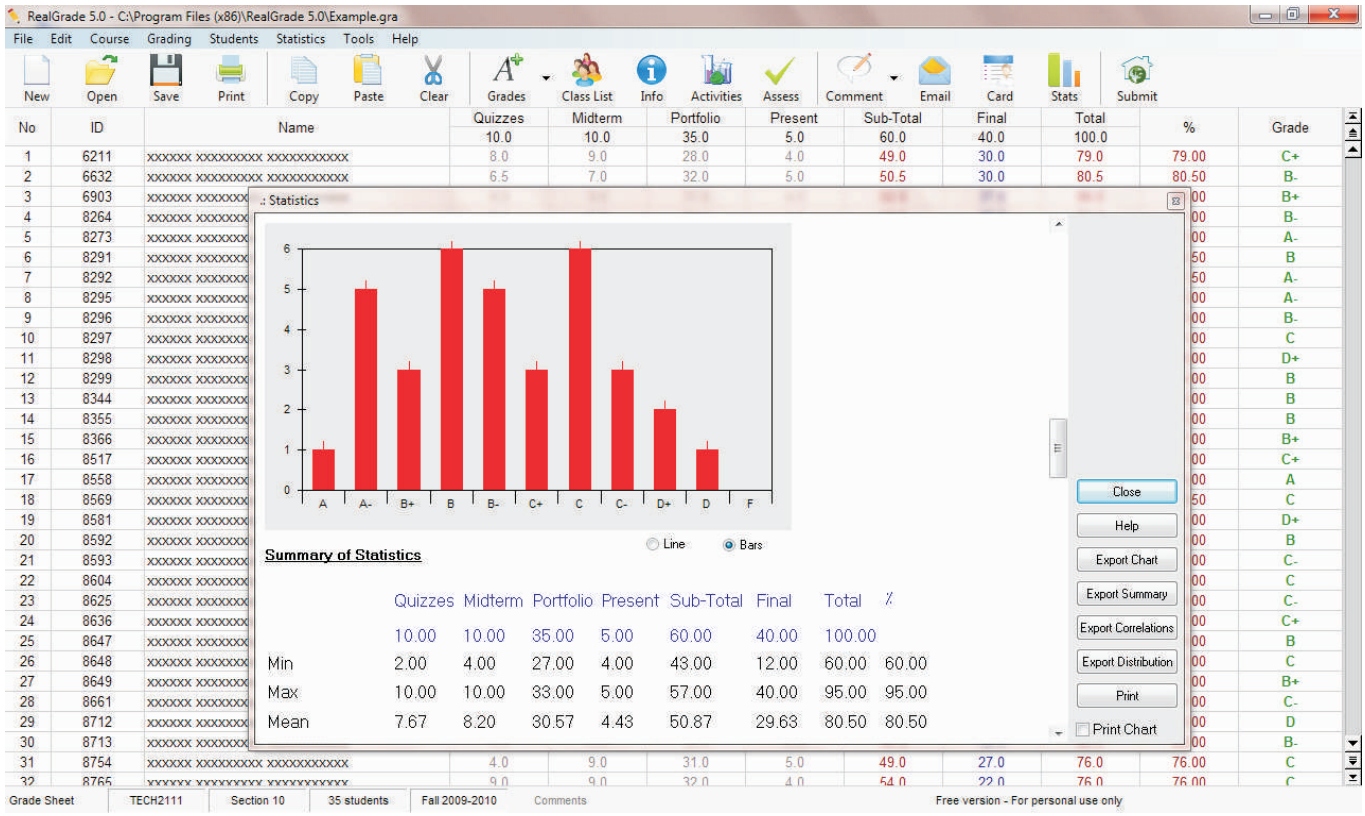

Figure 1. Real Grade Main User Interface

RealGrade during the Fall 2010 semester (January 2010). The only criterion for selection was that they indicated an interest in using RealGrade to grade and report student performance for the Fall 2010 semester and allowing the researcher to monitor their use and records intermittently.

One follow-up email invitation was sent. After this follow-up email message, responses were received from a total of 134 instructors (39.4%), who ranged widely in their specialization (e.g., education, engineering, commerce and economics, and arts and social sciences) and computer experience. None of the participants were required to have experience with specific computer applications. Many of them had a long and advanced experience in using spreadsheets to manage and report student grades. The majority had moderate experience in computer use. Participants were ensured that their identity and privacy would be protected during this study and that every attempt would be made to keep their students' personal data and grades confidential. At the end of the semester, only 116 instructors (86.6% of initial respondents) responded to the questionnaire.

Procedures

The first step in conducting a usability evaluation was adopting a framework with which to design the usability evaluation. Preece et al. (2002) describe the "DECIDE framework", which is suitable for software evaluation. It has six components: (i) determine the overall goals of evaluation; (ii) explore the specific questions to be answered; (iii) choose the evaluation paradigm and techniques to answer the questions; (iv) identify the practical issues that must be addressed, such as selecting participants; (v) decide how to deal with ethical issues; and (vi) evaluate, interpret, and present the data. This framework served as the basis for the design of the evaluation conducted in this study.

The evaluation paradigm for this study was usability testing. Participants' reactions were measured using the assessment tool usability questionnaire. Participants were then interviewed about their thoughts and were encouraged to add any other comments. Dillon (2001)

and Preece (1993) considered three different usability evaluation techniques that imply different types of evaluators, different numbers of users, and different types of data to be collected; these are user-based, expert-based and model-based evaluations. Preece argued that a user-based evaluation is the most realistic estimate of usability because it tests the extent to which the software supports the intended users in their work. Users are often asked to provide data on likes and dislikes through a questionnaire or interview. In this way, measures of usability can be derived and problems can be identified. However, because the main objective is to estimate the extent to which users in real situations can employ RealGrade effectively, efficiently and satisfactorily, the user-based technique was found to be the most useful for this study.

To determine the extent to which the software product is effective, efficient, and attractive to the participants under specified conditions, participants were asked to perform a series of specific tasks in RealGrade, each of which had several subtasks. These tasks included the following:

- Adding/importing a class list from the university Student Information System or Excel spreadsheet;
- Creating a new file, opening an existing file, and saving a current file;
- Adding, removing and sorting a student list;
- Defining and saving course-related information (e.g., semester, course title, instructor, and number of students);
- Linking content standards to academic activities (Figure 2);
- Assessing student performance based on performance standards;
- Categorizing, defining total points, and weighting academic activities (e.g., essays, quizzes, and tests);
- Commenting on students' performance and scores;
- Generating course statistics with different types of charts;
- Creating individual progress reports (grade card) for

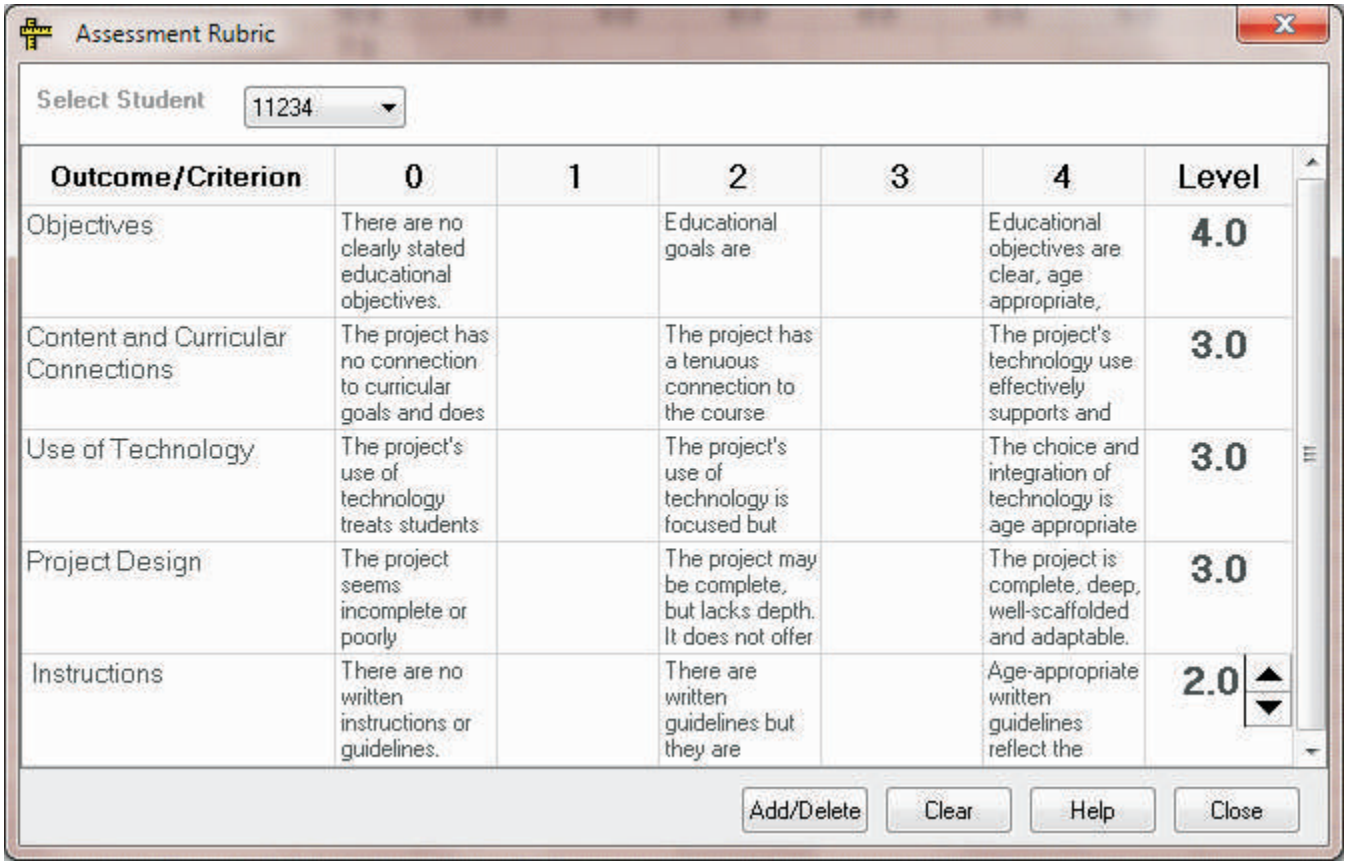

Figure 2. Standards-based Assessment Rubric

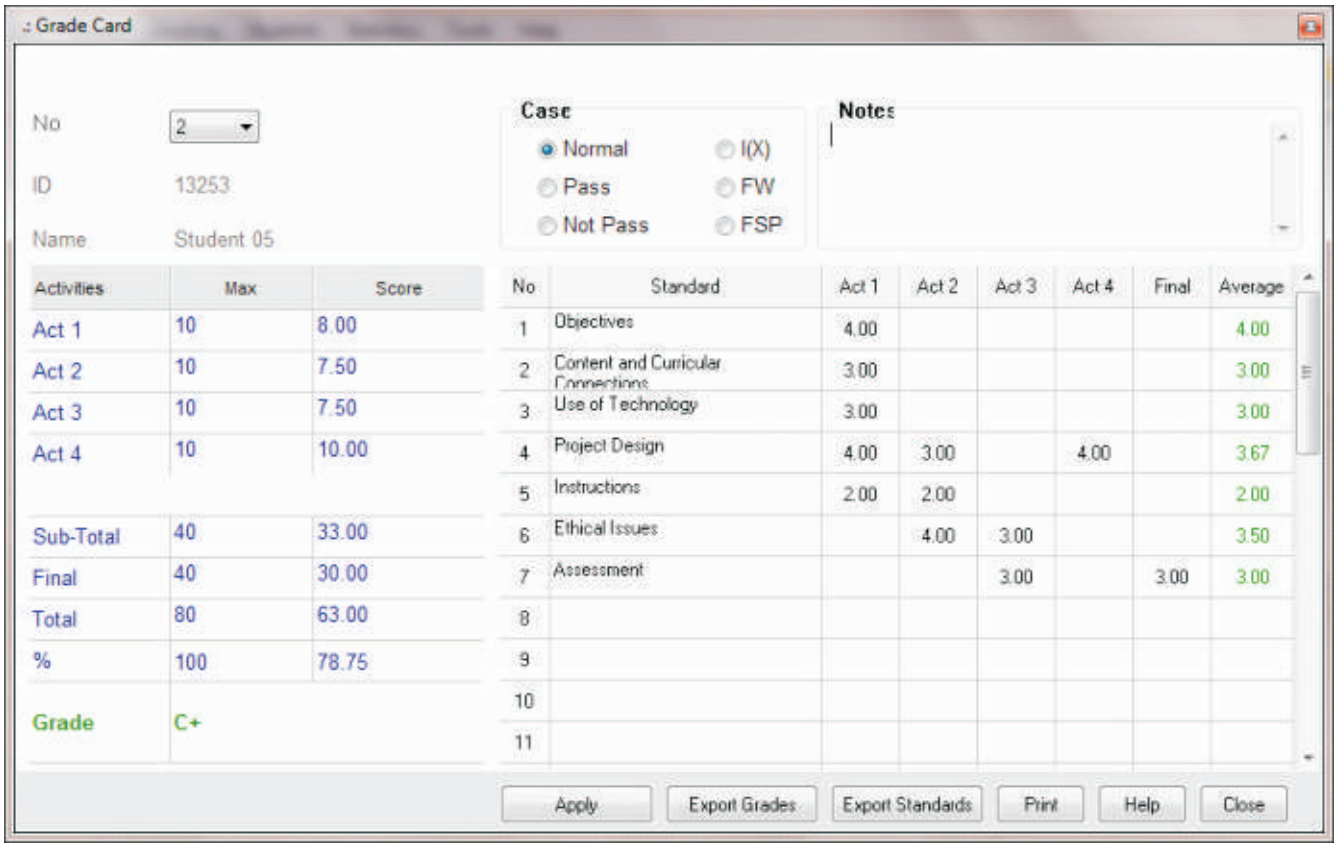

Figure 3. Standards-based Student's Grade Card

students (Figure 3);

- Sending assessment results to students via the email function;
- Printing out final grade sheets, attendance sheets, individual reports, course statistics; and
- Uploading grades to the University online grade entry system.

Participants were asked to record their experience and comments on each of these tasks and to record their thoughts regarding whether the software is useful, easy to use and appealing.

Instruments

Grading tool Usability Questionnaire (GUQ)

To answer the research questions, a usability evaluation questionnaire was developed in several phases using both quantitative and qualitative methods. The questionnaire development process occurred in four stages: delineation of relevant domains for the constructs of interest; questionnaire assembly and pilot testing; large-scale field-testing; and validation of instrument scores using factor analytic and correlation methods. The first step of instrument development involved reviewing the literature on software usability evaluation to conceptualize the domains that would directly influence usability of RealGrade. The review revealed many aspects that fell within three measures: usefulness of the software (effectiveness), ease of use (efficiency), and appeal (satisfaction) (Lohr, Javeri, Mahoney, Gall, Li, & Strongin, D., 2003).

Effectiveness is the main influence on the usability of computer software and is described as the perceived usefulness and importance of the software. Examples of items that could be used to measure perceived effectiveness include: "using the system in my job would increase my productivity" and "I would find the system useful in my job". Efficiency refers to the product's overall ease of use and simplicity. Responses to items such as "It would be easy for me to become skillful at using the

system" and "I would find the system easy to use" are used to evaluate efficiency of such software. Satisfaction measures how appealing the software is to users. Satisfaction is usually measured using items such as "I would like to use the software in my future career" and "I enjoy using the software".

The purpose of the second phase was to use the information in phase one to develop a multi-dimensional rating scale that could be used to assess the usability of RealGrade and to assess the content validity of its dimensions. Based on the conceptual definitions of the above measures of usability, each measure was examined for comprehensiveness. A pool of items was generated or modified to ensure appropriate and logical coverage.

A panel of six experts with adequate experience in usability testing, evaluation and measurement was enlisted to review and reflect on these measures and items. Three of the six panel members were instructional designers and three were SQU instructors. Panel members were tasked with suggesting items to add or delete and with commenting on each item's importance within each measure based on their understanding of the conceptual definition of the measure.

The resulting dimensions and items were pilot tested with a random sample of five instructors to assess the importance, clarity and wording of items. Items were revised based on the participants' degree of agreement and feedback. The revised dimensions were assembled into an online questionnaire. Instructors were asked to assess RealGrade using a Likert-style five-point rating scale ranging from "strongly agree" to "strongly disagree", in effectiveness sub-scale, or "very useful" to "not at all useful", and "very easy" to "not at all easy" in efficiency sub-scale (Table 1). In addition, open-ended questions probing positive and negative experiences were included to obtain any further suggestions or comments from the participants on each section. Example questions include the following: "what are the most/least useful features you found in RealGrade?", "what features/functions would you like to see added to the RealGrade?", "what are the features that saved your time and effort?" and "what are the features you most liked in RealGrade?"

In the third phase, the usability questionnaire with 48 items was field tested with a sample of SQU instructors. An invitation email message with information about the study and link to the online questionnaire was sent to the instructors at College of Education (N=140) asking them to download the software and complete the online questionnaire (January 2009). To maximize return rates, the Assistant Dean for Postgraduate Studies and Research sent an email requesting cooperation with the researcher. The response rate was monitored over a two-week period. After one follow-up email, responses were received from 32 instructors (23%). A Web site, which included background about RealGrade, purpose of the evaluation, and questionnaire instructions, was created with a link to the online questionnaire.

In the last phase, the psychometric characteristics of the questionnaire were investigated through the use of exploratory factor analyses and Cronbach's alpha. Because the questionnaire was divided into logically different sub-scales, common factor analysis was applied to verify whether the questionnaire measured only one dimension. Factors were extracted based on the proportion of variance explained by each factor. After list-wise deletion of the missing data, responses were available for 28 academics.

SPSS 13.0 was used to perform exploratory factor analysis. Principle component analysis with varimax rotation on the items identified three interpretable factors. Items with loadings greater than $\pm$ 0.40 were retained on the relevant factor, and items with loadings less than $\pm$ 0.40 were omitted.

Of the original 48 items included in these three factors, six were excluded from further analyses. A second factor analysis was then conducted on the remaining 42 items. The results showed that factor loadings ranged between 0.49 and 0.87 on the three measurements. These three factors accounted for 58.23% of the variance in the final version of the questionnaire. The eigenvalues of the three measurements from principle component analysis were each larger than 1:4.33, 2.74, and 2.03, respectively (Table 1).

These findings provide good evidence of content validity, as the highest factor loadings are central to the domains

No.	Item retained	Factor 1	Factor 2	Factor 3	Corrected item total correlation
	Effectiveness (=0.81)				
	Part 1: For each of the following items below, please tell us how much you agree or disagree with each statement. *(Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree)*				
1.	I would find using RealGrade useful in helping me to avoid errors often occur when performance standards are combined for assessment.	.78			.69
2.	RealGrade provides all the functions I need to assess students based on content standards.	.86			.83
3.	RealGrade provides efficient function to communicate standards-based assessment results with my students.	.81			.61
4.	RealGrade meets my standards-based assessment needs.	.77			.67
5.	Using RealGrade in grading would increase my productivity in standards-based reporting.	.69			.67
6.	Using RealGrade would improve my standards-based grading performance.	.70			.68
7.	I feel that there is a definite need for RealGrade in the college.	.78			.69
8.	RealGrade is a worthwhile standards-based grading and reporting tool.	.87			.81
9.	Using RealGrade would make it easier to submit students' grades.	.81			.61
	Part 2: When using RealGrade, how useful are the following options or functions? *(Very useful, Somewhat useful, Not very useful, Not at all useful, N/A)*	.79			.67
1.	Import class list directly from SIS or XLS.	.77			.74
2.	Comment on students' scores and activities.	.80			
3.	Track student attendance with absence warning indicators.				.77
4.	Define course content standards and attach them to each academic activity.	.78			.69
5.	Assess student achievement individually based on performance standards.	.86			.81
6.	Support different grading scales (under graduate, post graduate and custom).	.81			.61
7.	Standardize scores of assignments.	.77			.67
8.	Communicate standards-based grades with students via email.	.74			.68
9.	Generate course statistics for each student and the entire class.	.77			.69
10.	Generate individual standards-based grade reports.	.82			.70
11.	Attach students' documents/artifacts to their scores and assignments.	.78			.76
12.	Integrate content standards to the activities to provide rich information about student learning and course assessment.	.73			.68
13.	Submit grades directly to SIS.	.81			.67
	Efficiency (=0.52) For each of the following tasks below, please tell us how easy is RealGrade. *(Very easy, Somewhat easy, Not very easy, Not at all easy, N/A)*				
1.	Setup new class and course.		.60		.40
2.	Learn to use RealGrade for the first time.		.85		.35
3.	Find appropriate menus and dialogue boxes.		.60		.40
4.	Use the user's guide and instructions of use.		.67		.51
5.	Define and save course-related information.		.85		.54
6.	Correct and detecting entry errors in the spreadsheet.		.77		.42
7.	Link content standards to academic activities.		.60		.43
8.	Categorize, define total points, and weight academic activities.		.85		.44
9.	Create individual progress reports for students.		.77		.42
10.	Attach students' assignments to their grades.		.60		.43
11.	Assess student activities based on content standards.		.85		.52
12.	Print grade sheets and reports.		.77		.43
13.	Submit students' final grades.		.63		.44
	Satisfaction (=0.69) For each of the following items below, please tell us how much you agree or disagree with each statement. *(Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree)*				
1.	I feel comfortable when use RealGrade to assess my student performance based on course standards.			.49	.67
2.	I like the way that RealGrade uses to assess students based on performance standards.			.68	.63
3.	I would recommend RealGrade to my colleagues.			.70	.59
4.	RealGrade is an important tool for instructors.			.69	.56
5.	I feel I need to have RealGrade in my teaching.			.72	.54

Table 1. Retained items, rotated factor loading and eigenvalues for three factors of the GUQ (cont..)

No.	Item retained	Factor 1	Factor 2	Factor 3	Corrected item total correlation
6.	RealGrade is pleasant to use.			.65	.56
7.	Overall, I am satisfied with RealGrade.			.59	.57
	Eigenvalue	4.33	2.74	2.03	
	% of Variance	31.34	15.63	13.54	
	Total variance explained = 58.23%				
	Overall for the sub-scale = 0.76				

Table 1. Retained items, Rotated Factor Loading and Eigenvalues for Three Factors of the GUQ

assessed by questionnaire (Francis, Katz and Jones, 2000). Cronbach's coefficients for the three factors were 0.81, 0.52 and 0.69, respectively. Cronbach's for the entire sub-scale was 0.76. The item scale correlation coefficients range between 0.61 and 0.83 on the first factor, between 0.35 and 0.54 on the second factor, and between 0.54 and 0.67 on the third factor. These results confirm that the internal reliability index of the three constructs was adequate.

Based on logical and practical premises, the questionnaire was composed of three distinct constructs: effectiveness, efficiency, and satisfaction. For each construct, the mean response to the items was calculated, and the unit weighting of the items was used to construct factor score estimates. Relationships between constructs and entire scale were investigated (Table 2).

The inter-correlations show that, overall, each construct was significantly correlated with the other two constructs and with the entire scale. According to Harrison, Seeman and Behm (1991), this result provides further evidence for the consistency of the entire scale and for the convergent validity of each sub-scale. Therefore, it can be concluded that the three sub-scales and their constructs measure RealGrade usability in a coherent way.

Instructor interviews

For instructors to integrate the standards-based assessment tool into assessment and grading practices, they must view it in a positive manner, be comfortable with it and use it effectively. Therefore, determining what instructors were concerned with at the end of the implementation process was emphasized. A set of questions was asked to a group of instructors to gain a thorough understanding of the use of RealGrade and to provide rich detail and insights into instructors' experiences. Qualitative methods were used to provide consistent data. The purpose of these questions was to determine the perceptions of the instructors around usability issues of RealGrade. A series of semi-structured interviews were conducted after the implementation period. Individual interviews were conducted in person by the researcher.

Construct	Efficiency	Satisfaction	Scale
Effectiveness	.78*	.85*	.82*
Efficiency		.72*	.81*
Satisfaction			.79*

* Correlation is significant at the 0.01 level.

Table 2. Inter-correlation Matrix of Constructs

Data were analyzed to identify patterns, beliefs, values and practices as related to the instructors' RealGrade use. Instructor interview questions included the following:

- What is your overall impression of RealGrade?
- How useful do you find RealGrade in assessing students' learning based on performance standards?
- Do you think RealGrade positively or negatively influences your assessment activities?
- What obstacles are you facing in using RealGrade with your classes?
- If you could make one significant change to RealGrade, what change would you make?
- What features/functions would you like to see added to the new version of RealGrade?
- Would you recommend RealGrade to a colleague? Why?
- Do you have any other questions or comments about RealGrade or your experiences with it?

Although participants were prompted with questions, the main purpose was to get their subjective reactions toward

RealGrade. The total number of replies was counted and coded into three different aspects: effectiveness, efficiency, and satisfaction. Under each aspect, positive comments, negative comments, and suggestions for improving RealGrade were extracted.

Implementation

A three-stage methodology was adopted to implement RealGrade. The first stage included workshops and discussions focused on new trends in standards-based assessment, grading, and reporting, features and capabilities of RealGrade, the importance of RealGrade, and a tutorial about using RealGrade. These workshops were provided to instructors at their home colleges or departments. The second stage involved participant implementation of RealGrade during the Fall 2010 semester and performance of a series of specific tasks to grade student performance. The third stage involved consolidating participants' responses to the usability questionnaire to examine how usable RealGrade is. Issues highlighted from the implementation were explored further through interviews.

Results

Questionnaire Analysis

Using the Grading tool Usability Questionnaire (GUQ) and a series of individual interviews, usability was measured by the effectiveness, efficiency and satisfaction with which participants assess and report student performance using the tool. This section reports the results obtained from both the questionnaire and the interview.

Of the 116 instructors who participated in the study, a majority (79.9%) were male. Females made up 20.1% of the final sample (Table 3). Years of teaching experience ranged from 5 years to more than 15 years. The majority of instructors (57.8%) had 5-10 years of teaching experience. More than half of the participants came from the College of Education (56.7%). The rest of the participants came from a range of other colleges. Around two-thirds of respondents reported having moderate computer experience, and 20.2% indicated that they have good experience. Instructors were also asked to indicate how frequently they used RealGrade per week. The results indicated that the majority of respondents (65%) used RealGrade occasionally.

Demographics	%
Gender	
Male	79.9
Female	20.1
College	
Education	56.7
Arts & Social Sciences	17.5
Engineering	14.3
Commerce & Economics	11.5
Perceived computer experience	
Low	14.6
Moderate	65.2
High	20.2
Years of teaching experience	
Less than 5 years	24.6
5-10 years	57.8
More than 10 years	17.6
Frequency of RealGrade use	
Frequently	19.1
Occasionally	65.2
Seldom	15.7

Table 3. Instructor demographics and experience

The overall results showed that participants found RealGrade effective, efficient, and satisfactory (mean=4.22). In terms of effectiveness, participants strongly agreed or agreed that RealGrade is effective (3.93) in facilitating the process of standards-based assessment, has the tools needed to assess student performance and communicate grades with students, and increases their productivity. Participants also indicated that RealGrade is very useful or useful in importing class lists directly from SIS, and providing rich information regarding student performance (Table 4).

Participants were asked to indicate the most and least useful features they found. Only 26% of respondents completed this section, with 11% indicating more than one feature. Statements were coded and categorized, and they indicated that participants felt that using RealGrade can help instructors:

Scale/sub-scale	Possible range	Mean	Std. Deviation
Effectiveness	2.56-5	3.93	.6112
Efficiency	2.34-5	4.34	.4635
Satisfaction	2.21-5	4.19	.4521
Overall scale	2.48-5	4.22	.5356
Paired t tests for the means	Mean differences	S.D.	t
Effectiveness-efficiency	-.3300	.5347	-11.29*
Effectiveness-satisfaction	-.3500	.5439	-11.77
Efficiency-satisfaction	-.0023	.5667	1.87

** t is significant at the 0.001 level*

Table 4. Usability of RealGrade

- Provide individual report cards for student achievement;
- Submit final student grades and print grade sheets;
- Generate course statistics;
- Send e-mail reports;
- Comment on student scores and assignments;
- Grade each assignment based on one or more standard;
- Track student performance for each standard; and
- Sort students by name and grades.

When participants were asked about features or functions they would like to see added to RealGrade, they suggested many useful functions or features:

- "Provide a function that receives portfolio directly from students".
- "I want from the programmer who creates this useful program to receive student portfolio direct to RealGrade without using manual way".
- "Please include more activities i.e., activity 6, 7, 8 and the possibility to change values even if they are done".
- "I would like to see more of a combination for instructors who use the points possible system and an option to include a final exam. Also, increase more assignments under a category. I also believe maybe a tutorial video on the website would help with setup".
- "Weighted or unweighted assignments and assignment categories".

In terms of the efficiency, participants believe that RealGrade is an efficient tool (4.34). They indicated that learning to use RealGrade takes a short time, the user interface menus and dialogue boxes are favorable, assessing, and weighting, grading and managing student assignments based on performance standards are very easy tasks. Participants indicated many features they believe made RealGrade easy to use and saved their time:

- "Easy and familiar Windows interface".
- "Easy to understand. Each grading category lists all assignments and summary information".
- "Individual comments and attendance information".
- "Student individual card easy use and gives direct score to student".
- "Print a one student or entire class with a single click".
- "Class statistics can be viewed with a mouse click".
- "Performs all tasks involving grade calculation, averaging, and reporting, quickly and accurately".

In terms of satisfaction, the majority of participants reported that they liked and felt comfortable with RealGrade as a tool to assess student performance (4.19). They expressed that they would like to use it in the future and recommend it to their colleagues. In addition, participants reported that they were satisfied with many features and tools of RealGrade as follows:

- "I like you can have a view of all activities in the same screen".
- "I love the e-mail feature, toolbar and interface design".
- "I liked the way of changing grades, deleting an entire assignment, moving an assignment from one category to another, changing category weighting, curving grades for an assignment".

An examination of mean differences among the sub-scales shows that teachers scored highest on the efficiency sub-scale (an average of 4.34 per item) followed by the satisfaction sub-scale (4.19), and then the effectiveness sub-scale (3.93). The relatively lower score in the effectiveness sub-scale suggests that participants might not appreciate the usefulness of the standards-based assessment using the RealGrade (Table 4).

Results were further broken down by participants' computer experience and teaching experience. To investigate the relationship between participants' computer experience and their perceived usability, it should be mentioned that the literature has implicitly assumed a linear or logarithmic relationship between computer experience and perceived effectiveness, efficiency and satisfaction (Bozionelos, 2001). Therefore, types of relationships were examined throughout this study using scatter plots to determine what types of relationships existed. If a relationship seemed to be linear, the study continued to use that assumption. If it did not seem to be linear (e.g.,

logarithmic association), transformation of the scores was thought to be required.

Computer experience scores were categorized into three levels: low experience, moderate experience and high experience (Table 5). Plotting the initial results on a graph showed that they did fit a linear relationship, so no transformation was required. ANOVA tests were run to analyze the differences between computer experience groups and usability constructs. The results show that computer experience affected participants' satisfaction with RealGrade. Through a series of Scheffe tests (Post Hoc tests), it was found that participants having moderate or high computer experience tended to have statistically higher scores on the three sub-scales. In other words, participants with greater degrees of computer experience had higher perceptions of effectiveness, efficiency, and satisfaction with RealGrade.

Furthermore, the relationship between participants' teaching experience at SQU and their perceived usability was computed and categorized into three levels (Table 6). Plotting the initial results on a graph showed that they best fit a linear relationship. Therefore, an ANOVA and a series of Scheffe tests were used to analyze the differences. It was concluded that participants who had less than five years of teaching experience found RealGrade more effective, efficient and satisfactory than those who had more than five or ten years of experience.

Computer Experience	Effectiveness		Efficiency		Satisfaction	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
(1) Low	4.13	.5473	4.84	.4636	4.34	.5463
(2) Moderate	3.89	.6263	4.36	.5454	4.37	.4857
(3) High	3.78	.5465	4.45	.3781	4.45	.5769
F	4.52*		8.47**		2.11*	
Scheffe test	(3)>(2) (3)>(2)>(1)		(3)>(2)>(1)		(3)>(2)>(1)	

* F is significant at the 0.01 level
** F is significant at the 0.001 level

Table 5. ANOVA Results of Usability by Computer Experience

Teaching experience	Effectiveness		Efficiency		Satisfaction	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
(1) Less than 5 years	4.03	.5018	4.45	.4174	4.22	.5594
(2) 5-10 years	3.94	.6175	4.27	.5226	4.27	.4838
(3) More than 10 years	3.68	.7645	4.11	.3945	4.11	.5647
F	5.61*		9.36*		2.03*	
Scheffe test	(1)>(2)		(1)>(2)>(3)		(1)>(2)>(3) (2)>(3)	

* F is significant at the 0.01 level

Table 6. ANOVA Results of Usability by Teaching Experience

Interview Analysis

To learn more about the impression of instructors regarding the usability of RealGrade and to validate results after the usability survey, eight participants, representing the four colleges were randomly selected (7% of the total number of participants) according to the percentage of participants from colleges, as represented earlier in Table 3. The responses to the eight interview questions are organized, analyzed, and coded to address the research questions. However, since many responses contained multiple beliefs, the number of codes assigned to each passage varied. Responses are categorized according to the first three research questions and the type of feedback (general or distinctive) as shown in Table 7.

Overall, feedback from interviewees showed that participants found RealGrade useful and easy to use. Responses also indicated that participants felt RealGrade was a satisfactory way to record student assessment information and to conduct standards-based assessment. Participants indicated recognizing the usefulness and ease

Results	General patterns (frequency)	Distinctive viewpoints (frequency)
1. Effectiveness (usefulness)	1.1. Avoid grading errors (2) 1.2. Error-free and accurate (2) 1.3. Helpful in grading and reporting (2) 1.4. Important for instructors (1) 1.5. Improve instructor performance (1) 1.6. Increase instructor productivity (2) 1.7. Multiple output formats (1) 1.8. Save time (3) 1.9. Simplify standard-based assessment (1) 1.10. Useful (4)	- Customized grading scale (1) - Professional way fir grading (2) - Use well-defined rubric (1) - Collaborative (1)
2. Efficiency (ease of use)	2.1. Appropriate documentation (2) 2.2. Compatible with my system (4) 2.3. Easy to learn (5) 2.4. Easy to setup (3) 2.5. Flexible (2) 2.6. Simple and attractive (3) 2.7. User-friendly (2)	- Keep students updated about their performance (1) - Essential for quality assurance and accreditation (2) - Easier than MS Excel grade sheets (1) - Mobility (1)
3. Satisfaction	3.1. Like to use (5) 3.2. Interesting (3)	- Recommended for official use (4)

Table 7. Analysis of Interview Results (N=7)

of use of RealGrade in assessment from both a practical and educational perspective(*).

For example, in terms of effectiveness (usefulness) of the program, an instructor expressed the following:

> *"Thank you very much for this nice program. I found the program to be a very useful tool for me as an instructor. The good thing about it is that it is easy to understand and it has many good features which are important in grading".*

Similarly, another instructor commented that RealGrade is important enough to be part of the college assessment practices.

> *"It is a wonderful program that saves my time when grading. It should be a major priority of the College to integrate this standards-based gradebook into the college".*

In addition, an instructor expressed that the structure of RealGrade was clear and understandable and reduces the need for additional record software:

> *"When I used RealGrade for the first time, I liked its simple interface and found its look pleasant. It was easy to learn and its contents are organized in a proper way. I found it more effective than spreadsheets and databases. I found most of features as important for standards-based assessment work and I do not need spreadsheets of marks any more".*

One instructor suggested that RealGrade needs more collaborative features to encourage sharing of grades and student performance among instructors.

> *"It is very helpful to share assessment results with my students, but it should provide some extra features for collaborative work with colleagues. It should provide a real time feature".*

A further benefit of RealGrade is the way coursework and assessment could be processed quickly and efficiently. Two instructors reported that:

> *"RealGrade considerably improves the assessment and feedback process. It improves the management of coursework and feedback. Grading time is shorter and the feedback is sent to students faster than before", and "I am very pleased and impressed that you have taken the time to develop this useful program. I used to hate preparing grade reports manually or using Excel. It took so much time before and after the final exams. Now I just enter assignments, scores, and comments, and click a button to print or send may grades directly to the SIS".*

In terms of standards-based assessment, an instructor expressed that

> *"My overall impression is very positive. It is very useful and all instructors at SQU should use it. The most important feature of RealGrade over Excel spreadsheet is that RealGrade includes the ability to make standards-based assessment meaningful, and to create, print, or e-mail detailed individual reports to students. It is easier to setup a class using standards than other programs".*

In addition, participants reported many advantages of RealGrade, such as the ability to import and present student information quickly, categorize and weight class assignments accurately, link one or more standards to student assignments, and automatically calculate standards-based grades.

In terms of efficiency (ease of use), one instructor who had a long experience in using his own Excel spreadsheets for standards-based assessment did not feel that RealGrade could, or should, replace Excel spreadsheets. He argued the following:

> *"What I really reject about RealGrade is anything that is already done on Excel".*

In addition, one instructor denied that he was generally unwilling to use standards-based grading in assessment in his course and believed that more training is required to learn how to use RealGrade.

> *"Standards-based grading using RealGrade is a difficult job to acquire, but I have to use it in ways that would align with the university philosophy. I need more training on using the software".*

This idea was justified by the point of view of another instructor, who emphasized the effect of the instructor's background on the success of RealGrade. The instructor

(*) Most interviewees' responses were translated from Arabic by the interviewer.

stated the following.

> *"Learner analysis grade card is really of critical importance. One needs to consider the background and knowledge of the instructor who is doing standards-based assessment using the gradebook. Whether the instructor is familiar with computers, what is his/her experience in assessing student performance; there are many factors".*

From a technical perspective, an instructor proposed an interesting idea for developing grading tool. He stated that there is a need for a portable version of RealGrade:

> *"An advantage of running RealGrade software from a removable USB drive is that I do not need to install the software at home, computer lab, office, etc and I can access my students grades everywhere."*

Lastly, all of the participants complained about the time limitation. They all stated that the 15-week duration of the semester, and therefore the study implementation period, was not sufficient to develop their course and performance standards and integrate RealGrade into their assessments. They also stated that the workloads and duties of teaching should be well planned to be able to manage and assess student performance in the most efficient and effective way possible.

Discussion and Conclusion

Universities have a professional responsibility to ensure that their programs and graduates are of the highest quality. Meeting this responsibility requires incorporating content and performance standards into the university curriculum programs and assessments. However, because instructors must provide evidence that students completing their degrees have performed at acceptable levels, the need has emerged to develop a standards-based assessment tool that allows for more accurate and relevant grading and reporting as well as tracking of content standards.

To meet this need, this study designed a assessment tool to judge and grade student performance against a set of course standards using a rating scale based on explicit rubrics. This assessment tool provided useful tracking and reporting features to instructors. These features served to facilitate and promote a greater understanding of student performance. The development of the preliminary version of RealGrade allowed usability of that prototype to be investigated.

The usability questionnaire and individual interviews provided useful feedback regarding the usability of RealGrade. Regardless of their previous assumptions, educational philosophy, technical skills, and level of teaching experience, participants' responses overall were extremely positive. Instructors agreed that RealGrade was useful and an easy-to-use tool that facilitated the process of gathering and judging grades to decide whether students achieved content standards. RealGrade assisted instructors in communicating this information regarding student performance, which is the primary purpose of grades. More positive comments than negative comments were provided in individual and group interviews. The majority of instructors stated that RealGrade was useful, easy to use, and appealing because it simplified standards-based assessment and provided a wide range of options to communicate grades.

Davis (1989) showed that perceived usefulness and ease of use are each highly correlated with self-reported use and future use. Ease of use appears to be a causal antecedent of usefulness, with little direct effect on use. In addition, Igbaria, Zinatelli, Cragg and Cavaye (1997) noted the importance of ease of use, or complexity, in the decision to use software. Specifically, it has been shown that the complexity of the innovation has a significant negative relationship with adoption of the new application. Rogers (1995) found that the relative advantage (usefulness), as perceived by the users, is positively related to the innovation's rate of adoption. He discusses some forms of incentives that may provide that relative advantage mentioned. Huff and McNaughton (1991) found that while the users perceived the usefulness of the software, the benefits of using the system needed to be communicated further to the users. According to the Technology Acceptance Model (TAM), perceived ease of use and perceived usefulness mediate all other external variables that are likely to influence adoption and usage decisions by the individual (Mathieson, 1991). In other words, people are more likely to use software they perceive

as easy to use and useful for performing job tasks.

Further investigation of variables that affected instructors' perceived usability indicates that differences among instructors are vital to the eventual acceptance and implementation of the standards-based assessment tool. This result is consistent with previous findings concerning the impact of instructors' computer experience on computer use. For example, Sadik (2006) found that prior computer knowledge and experience influence the acceptance and use of new computer systems. In addition, the number of years of teaching experience of an instructor has shown a high level of significance in all of the three aspects of the usability questionnaire. This also appears to be consistent with findings from previous studies. Henry and Stone (1997) linked years of teaching experience with teacher age, stating that typically teachers with more years of experience tend to have more trouble with the integration of technology. Therefore, if University officials attempt to integrate RealGrade, a major priority should be knowing who they are asking to use the new tool. They should consider degrees of teaching experience, computer experience and training received in standards-based assessment approaches.

Lee, Kim, and Lee (1995) looked at the role of training in user acceptance of new technology. They asserted that proper training can alleviate individual differences while increasing job satisfaction, information system satisfaction and acceptance, end-user ability, and system utilization. With proper training for instructors who have little computer experience and full utilization of RealGrade, the effort and time instructors spend in judging and calculating grades can be reduced.

However, although RealGrade simplified the assessment process and allowed instructors to summarize data on student performance, it is the instructor's responsibility to ensure consistency in the evidence gathered, decide what information goes into the calculation, and define what weight should be assigned to each activity to generate the most accurate and fairest description of each student's achievement and level of performance.

The possibilities for future research are exciting. With respect to software development, RealGrade should be considered a work in progress. Feedback from instructors can provide indicators regarding how and what features are desirable. For example, further research is needed to develop RealGrade report card to help instructors and students understand the standards-based assessment information included and to make it more comprehensible. In addition, new features need to be added to allow instructors and students to distinguish the difference between formative evidence, which has the purpose of examining student understanding and guiding instructional revisions, and summative evidence, which is gathered to determine a final grade.

References

[1]. Bevan, N. (2001). International standards for HCI and usability. *International Journal of Human-Computer Studies*, 55, 533-552.

[2]. Bozionelos, N. (2001). Computer anxiety: Relationship with computer experience and prevalence. *Computers in Human Behavior*, 17(2), 213-224.

[3]. Burger, D. (1998). Designing a Sustainable Standards-based Assessment System. Mid-continent Regional Educational Laboratory, Inc.

[4]. Daintith, J. (2004). Software prototyping. *A Dictionary of Computing*. Retrieved from http://www. encyclopedia. com/doc/1O11-softwareprototyping.html

[5]. Davis, D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, *MIS Quarterly,* 13(3), 318-340.

[6]. Dillon, A. (2001). Usability evaluation. In W. Karwowski (ed.) *Encyclopedia of Human Factors and Ergonomics,* London: Taylor and Francis.

[7]. Ebel, L. & frisbie, a. (1986). Essentials of educational measurement (4th ed.). *Englewood Cliffs*, N.J.: Prentice-Hall.

[8]. Francis, L., Katz, Y. and Jones, S. (2000). The reliability and validity of the Hebrew version of the Computer Attitude Scale. *Computers and Education*, 35(2), 149-159.

[9]. Guskey, T. (2001). Helping Standards Make the Grade. *Making Standards Work,* 59(1), 20-27.

[10]. Harrison, P., Seeman, J., & Behm, R. (1991). Development of a distance education assessment

instrument. *Educational Technology Research and Development,* 39(4), 65-77.

[11]. Henry, W. & Stone, W. (1997). The development and validation of computer self-efficacy and outcome expectancy scales in a non-volitional context. *Behavior Research Methods, Instruments, & computers.* 29(4), 519-527.

[12]. Hills, R. (1991). Apathy concerning grading and testing. *PhiDelta Kappan*, 72(7), 540-545.

[13]. Huff, L. & McNaughton, J. (1991). Diffusion of an Information Technology Innovation. You and the Computer, *Summer*, 25-30.

[14]. Igbaria, M., Zinatelli, N., Cragg, P. & Cavaye, A. (1997). Personal Computing Acceptance Factors in Small Firms: A Structural Equation Model. *MIS Quarterly*, 21(3), 279-305.

[15]. ISO 9241 (1998). Ergonomic requirements for office work with visual display terminals. Part 11: *Guidance on usability*. Retrieved from http://www.iso.org /iso/ catalogue _ detail. htm?csnumber=16883.

[16]. Kennedy, D. (1998). Software Development Teams in Higher Education: an Educator's View. In R. M. Corderoy (ed.) FlexibilITy: The next wave. ASCILITE 98. *Proceedings of the 15th Annual Conference of the Australian Society for Computers in Learning in Tertiary Education*. Australia, University of Wollongong, 373-386

[17]. Khattri, N., Kane, M., & Reeve A. (1995). How performance assessments affect teaching and learning. *Educational Leadership*, 53, 80-83.

[18]. Lee, M., Kim, Y. & Lee, J. (1995). An Empirical Study of the Relationships among End-User Information Systems Acceptance, Training and Effectiveness. *Journal of Management Information Systems*, 12(2), 189-202.

[19]. Lohr, L. (2000). Three perceptual principles for instructional interface design. *Educational Technology*, 40(1), 45–52.

[20]. Lohr, L., Javeri, M., Mahoney, C., Gall, J., Li, K. & Strongin, D. (2003). Using Rapid Application Development to Improve the Usability of a Preservice Teacher Technology Course. *ETR&D*, 51(2), 41–55.

[21]. MacCormack, A. (2001). Product-Development Practices that Work: How Internet Companies Build Software, Sloan Management Review, *Winter,* 75-84.

[22]. Mathieson, K. (1991). Predicting User Intentions: Comparing the Technology Acceptance Model with the Theory of Planned Behavior. *Information Systems Research*, September, 173-191.

[23]. Nielsen, J. (2003). *Usability 101: Fundamentals and definitions - what, why, how?* Retrieved from http://www. useit.com/alertbox/20030825.htm

[24]. Officeware Corporation (1998). *A Structured Approach to Programming Modern Business Applications: Methodology Definition, Whitepaper 1*, Retrieved from http://www. oware. com/Subpages/radmethodology.htm

[25]. Oosterhof, C. (1994). *Classroom applications of educational measurement (2nd ed.)*. New York: MacMillan.

[26]. Preece, J. (Ed.) (1993) A Guide to Usability: Human Factors in Computing. Wokingham, UK: Addison-Wesley.

[27]. Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design: Beyond human-computer interaction.* Hoboken, NJ: John Wiley & Sons.

[28]. Reed, D. (1996). High School Teachers' Grading Practices: *A Description of Methods for Collection and Aggregation of Grade Information in Three Schools,* Unpublished Ph.D. Dissertation, University of California.

[29]. Roblyer, D., Edwards, J., and Havrileck, A. (1999). *Integrating Educational Technology into teaching.* Upper Saddle River, NJ & Columbus, OH: Merrill, an imprint of Prentice Hall.

[30]. Rogers, M. (1995). *Diffusion of Innovation*, 4th ed., New York: The Free Press.

[31]. Rozanski, P., and Haake, R. (2003). Curriculum and content: The many facets of HCI. Paper presented at the *4th Conference on Information Technology Curriculum on Information Technology Education,* Lafayette, Indiana, USA.

[32]. Rubin, J. (1994). *Handbook of Usability Testing*. New York: John Wiley and Sons.

[33]. Rushby, J. (1997). Quality criteria for multimedia. *Association for Learning Technology Journal*, 5(2), 18-30.

[34]. Sadik, A. (2006). Factors Influencing Teachers'

Attitudes Toward Personal Use and School Use of Computers: New Evidence From a Developing Nation. *Evaluation Review*, 30(1), 86-113.

[35]. Scacchi, W. (2000). Understanding Software Process Redesign using Modeling, Analysis and Simulation. *Software Process Improvement and Practice*, 5(2), 183-195.

[36]. Scacchi, W. (2001). Process Models in Software Engineering. In Marciniak, J. (ed.), *Encyclopedia of Software Engineering*, 2nd Edition, New York: John Wiley and Sons, Inc.

[37]. Walvoord, E. (2004). *Assessment Clear and Simple: A Practical Guide for Institutions*, Departments, and General Education. San Francisco: Jossey-Bass.

[38]. Wesley. Pressman, S. (1997). *Software engineering: A practitioner's approach*. (Fourth ed.). New York: McGraw-Hill Companies Inc.

ABOUT THE AUTHOR

Alaa M. Sadik received his Ph.D. degree in Educational Technology from the Institute for Learning, the University of Hull, United Kingdom, in 2003. He works as an Associate Professor of Instructional Technology at South Valley University, Egypt and currently he works at Sultan Qaboos University, Oman.. He is interested in the integration of computer and Internet applications in education to improve teaching and learning particularly in developing countries.