


Alternative Assessment Approaches and Quality Product Design Within Web-Based Learning Environments

Ahlam Mohammed Al-Abdullatif, King Faisal University, Saudi Arabia

 <https://orcid.org/0000-0003-2815-1137>

ABSTRACT

The aim of the present study is to identify the impact of the alternative web-based self and peer assessment approaches on improving the quality of student educational projects. In this context, a study was carried out during the second semester of the 2017-2018 academic year among 48 postgraduate students at King Faisal University. Results indicated that both self and peer-assessment approaches are effective when assessing the quality of educational products. The results also showed that the extent of student experience with the self-assessment approach affects their assessment credibility and objectivity. This study emphasized the value of utilizing alternative assessment approaches in web-based learning environments as means of improving student performance, particularly when designing educational products. It may have theoretical and pedagogical implications for learners and teachers.

KEYWORDS

Alternative Assessment, Peer-Assessment, Quality Product, Self-Assessment, Student Performance, Web-Based Learning

INTRODUCTION

The assessment of student learning is one of the most important aspects of the educational process. It helps those who are involved in learning and teaching judge the level of achievement of learning outcomes and their appropriateness to student skills and capabilities. The traditional assessment system, derived from the perspective of the teacher only, does not help to measure all aspects of student learning. To address the problems of the traditional system, alternative assessment approaches have emerged. These promote the active participation of students in the assessment process and, by increasing their awareness of the assessment criteria, improve and develop the quality of learning (Gibbs, 2010; Price, Carroll, O'Donovan, & Rust, 2011; McKeivitt, 2016). The present study aims to assess the improvement of student learning performance using two types of alternative assessment approaches: self-assessment and peer-assessment.

DOI: 10.4018/IJWLTT.20200701.oa1

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

LITERATURE REVIEW

Alternative Assessment Approaches

Many researchers have recognized that alternative assessment techniques are effective and influential means of assessing learner educational development. Alternative assessments have been found to facilitate the process of instruction, as they include procedures and techniques that are easily incorporated into student activities (Iraji, Enayat, & Momeni, 2016).

The alternative assessment techniques of self and peer-assessment foster student learning and develop a number of essential learning skills, such as reflection, critical thinking, and self-awareness. They also help to raise the awareness of learners about the assessment process. As Ashraf and Mahdinezhad observed, “Peer and self-assessment, where students assess each other and themselves, can encourage them to take greater responsibility for their learning, for example, by encouraging engagement with assessment criteria and reflection of their own performance and that of their peers” (2015, p. 111).

Self-Assessment

Self-assessment can be defined as:

a process of formative assessment during which students reflect on and evaluate the quality of their work and their learning, judge the degree to which they reflect explicitly stated goals or criteria, identify strengths and weaknesses in their work, and revise accordingly. (Andrade & Du, 2007, p. 160)

Several characteristics of the self-assessment approach are emphasized in the literature: desired learning outcomes are explicated to students through a rubric and/or model examples of student work; students assess their draft work according to the rubric and/or the examples; and students use the feedback generated by their teacher to guide them in improving the quality of their work (Price et al., 2011; Khonbi & Sadeghi, 2013; Põldoja, Väljataga, Laanpere, & Tammets, 2014; Zarei & Usefli, 2015). Self-assessment is a fundamental component of effective learning. Encouraging student autonomy, independence, and self-regulation, it encourages learners to be proactive, influencing their own learning rather than waiting for others to do so (Boud, 2013). As Boud (2013) puts it, “Learning can only be effectively undertaken when the learner monitors what is known, what remains to be known and what is needed to bridge the gap between the two” (p. 15). Educational psychologists include this kind of self-monitoring as part of the metacognitive skill set, which is now central in cognitive theories of learning (Biggs & Moore, 1993). Self-assessment is criterion-referenced; thus, it requires students to learn how to identify the characteristics of a strong piece of work before applying this information to their own work (Andrade & Du, 2007; Boud, 2013). Andrade and Boulay (2003) have suggested that self-assessment involves ongoing monitoring and regulating one’s thinking processes and task performance. This monitoring ultimately enhances the quality of learning. Additionally, researchers have emphasized the power of feedback in self-assessment when it comes to guiding student efforts and promoting their learning strategies (Sung, Chang, Chiou, & Hou, 2005; Andrade & Du, 2007; Butler & Lee, 2010; Boud, 2013; Birjandi & Bolgori, 2015).

Although self-assessment approach encourages engagement and responsibility among students by leading them to reflect on their own work and to judge the quality of their performance in terms of given assessment criteria, it has been critiqued in the literature because of the risk of bias, students might tend to assess themselves subjectively. This bias could result in unreliable assessments: “Self-evaluation has a risk of being perceived as a process of presenting inflated grades and being unreliable; and students feel ill-equipped to undertake the assessment” (Ashraf & Mahdinezhad, 2015, p. 111).

Peer-Assessment

Peer-assessment is defined as a process whereby students rate their peers and make judgments about other student's work. It is a tool meant to provide peers with constructive feedback; in this approach, students are involved both in the learning and in the assessment process (Sitthiworachart & Joy, 2003). The peer-assessment approach has become popular in contemporary educational contexts.

Peer-assessment has been advocated for its many characteristics that enhance student learning. Liu and Carless (2006) have emphasized the power of feedback as a learning element in peer-assessment, noting its great potential for enhancing student learning. Peer-assessment enables students to take an active role in managing their own studies, which is an essential component of self-regulated learning (Butler & Winne, 1995). In this process, students monitor their work using internal and external feedback as facilitators. Falchikov (2001) has found evidence that peer feedback enhances student learning because learners actively engage in developing subject-matter understanding. Meanwhile, Liu and Carless (2006) have indicated that it is important to engage learners with peer-assessment in order to allow students to make their work public while creating conditions that facilitate social learning: "In order to clarify notions of quality, learners need to analyze real, illustrative exemplars. This is where examining the work of peers offers meaningful opportunities for articulating discipline-specific knowledge, as well as criteria and standards" (p. 281). For their part, Birjandi and Bolgori (2015) suggest that peer-assessment might increase reflection and generalization to new situations, thus promoting metacognitive self-awareness and a sense of ownership.

Additionally, peer-assessment has been described as an alternative assessment method based on a constructivist approach that presents several advantages for learning. These include improving the reliability of assessing learner performance; increasing student responsibility for self-reflection; raising learner motivation; emphasizing evaluation and discussion skills; involving relevant feedback from multiple raters (peers as well as the instructor); and, by perceiving mistakes as opportunities rather than failures, enhancing student performance (Price et al., 2011; Kao, 2013; Uto & Ueno, 2016). Like self-assessment, peer-assessment also has certain disadvantages, including the tendency of students to award everyone the same mark and their reluctance to make judgments regarding their peers (Ashraf & Mahdinezhad, 2015).

Alternative Assessments in Web-Based Learning Environments

Web-based learning environments are technological innovations that have imposed themselves on the educational system. Web-based learning environments include many features and possibilities that address learner needs and requirements. Even at a distance, these technologies offer learning tools that meet diverse needs and styles of learning through multiple sources and varied materials and tools. Importantly, web-based learning environments are learner-centered environments that enable learners to actively engage in self and peer-assessment process through different tools and techniques. Features, such as synchronous and asynchronous learning activities, provide means of effective communication while reinforcing teaching and learning processes (Atan, Rahman, & Idrus, 2004). Web-based learning environments hold great promise in creating collaborative learning opportunities for learners, allowing them to share their opinions and interact as they would in a single classroom working together to build their own learning; in essence a digital environment can enrich opportunities for learning (Hsu, Ching, & Grabowski, 2014).

Studies recognize self and peer-assessment as two activities beneficial to the temporary learner (Elliott, 2008). Indeed, they are considered important techniques in advancing student learning in web-based learning environment (Sung et al., 2005; Liu & Carless, 2006; Andrade & Du, 2007; Butler & Lee, 2010; Boud, 2013; Birjandi & Bolgori, 2015). Further, students highlighted the value of these assessment techniques in terms of the guidance provided by the rubric prior to feedback and the fact that feedback was specific about how they could improve. These findings support those of other studies in the related literature of self and peer-assessments (Vickerman, 2009; Leach, 2012; Liu & Lee, 2013; Ashraf & Mahdinezhad, 2015; Zarei & Usefli, 2015; Irajii et al., 2016).

In terms of web-based learning environment, several studies in the literature have demonstrated the effectiveness of using web-based self and peer-assessment methods in enhancing student learning and performance. Sung et al. (2005) have found that, upon completion of a web-based peer and self-assessment system, students demonstrated a better quality of their work. Another study by McKeivitt (2016) investigated the impact of web-based assessment tools on third-year humanities students in terms of student engagement with the lesson criteria, self-assessment, and feedback. The findings indicated a significant improvement in student performance between the first draft and final submission of their works. Additionally, a study by Tas, Cetinkaya, Karakaya, and Apayadin (2013) showed that students tested with a web-based measurement tool demonstrated a statistically significant higher achievement in a science unit. Similarly, Domínguez, Jaime, Sánchez, Blanco, and Heras (2016) conducted a comparative study in a web-based environment measuring the consistency and differences between self-, peer and teacher-assessment; this research showed that alternative assessment experiences correlated with learning gains.

Another study by Andrade and Du (2007) suggests that students had positive attitudes toward self-assessment as they felt that they could effectively self-assess when they knew the teacher's expectations; these students claimed to have used self-assessment as a way of checking their work and guiding revision. The students pointed out several benefits of self-assessment, including improved grades, higher quality of work, and added motivation and learning.

Some studies have focused on determining the consistency and differences between self-, peer and teacher-assessment. For example, Chang, Tseng, and Lou (2012) conducted a comparative study to explore the consistency and differences of web-based portfolio assessment, finding high level of consistency between self- and teacher-assessment, but low level of consistency between peer and self as well as teacher-assessment. However, Domínguez et al. (2016) in their study of using web-based self and peer-assessment techniques found a high level of consistency among self, peer, and teacher-assessment.

In addition, several studies have noted the reliability and objectivity of both self and peer-assessment scores among students using web-based interactive software. (Xiao & Lucking, 2008; Chen & Tsai, 2009; De Wever, Van Keer, Schellens, & Valcke, 2011; Põldoja et al., 2014; Uto & Ueno, 2016). The studies varied in their determination of which method was the more effective in the assessment process. Some of the studies showed that self-assessment of learner performance was more effective than either peer-assessment or teacher-assessment (Esfandiari & Myford, 2013; Zarei & Usefli, 2015), while other studies found the contrary (Khonbi & Sadeghi, 2013; Zarei & Sayar Mahdavi, 2014; Ashraf & Mahdinezhad, 2015). In Leach's study (2012), no difference was found between the effectiveness of student self-assessment and teacher-assessment. However, Chang et al. (2012) indicated that both self and peer-assessment are considered to be valuable learning activities that are essential for evaluating learning as a 'multi-dimensional' process, thus increasing assessment validity and reliability (p. 312).

RESEARCH CONTEXT AND OBJECTIVE

Current learning approaches emphasize the active engagement of students in their own learning as well as learner responsibility; these methods uphold a dialogical, collaborative model in teaching and learning contexts. In conventional classroom assessment processes, the teacher plays the dominant role and tends to retain all ownership and power (Spiller, 2012). The recently-developed idea of web-based assessment has been widely applied in the field of education to enhance instructor professional development while promoting the acquisition of relevant skills for both teachers and students (Kilimci, 2010; Liu, Shih, & Tsai, 2011; Liu & Lee, 2013; Cheng, Liang, & Tsai, 2015; Uto & Ueno, 2016).

In the context of Saudi Arabia's tertiary education, many institutions have begun adopting web-based learning environments in constructing and delivering teaching and learning. However, there is a vast number of students enrolled in colleges, causing a major teaching load on instructors. As a result,

use of the Blackboard Learning Management System (BLMS) has remained traditional, limited to uploading course materials and learning resources (Al-Abdullatif, 2012). It is uncommon for faculty members to use Blackboard's advanced features and tools for assessment, particularly those with little technological savvy. Therefore, the majority of instructors relies heavily on traditional assessment methods, such as quizzes, assignments, and final exams. Alternative web-based assessment approaches, particularly self and peer-assessment, are not typically used in the Saudi education system. Many doubt their objectivity and reliability, preferring to defer to the teacher's assessment. Even if these methods are implemented, they usually are not considered in the students' final grades. Therefore, the present study aimed to explore and identify the impact of alternative web-based assessment approaches in self and peer-assessment on improving the quality of student design for their educational product presented, in this context, as a WebQuest projects. The results will help teachers to determine whether or not self and peer-assessment are suitable methods for assessing the quality of student educational products; this study will also help to determine the objectivity and reliability of these methods.

Quality Product Design

Assessing the quality of product design involves comparing student performance or output to particular criteria by means of a rubric. According to Brophy, "A rubric is a measurement tool that describes the criteria against which a performance, behavior, or product is compared and measured" (2012, p. 3). For this study, an analytic rubric was developed which described the level of achievement associated with the lesson criteria by providing a separate score for each. Such rubrics are often developed in the form of a matrix that consists of four basic parts (University of Hawaii, 2017):

1. **A task description:** The learning outcome being assessed, or instructions students received for an assignment;
2. **The characteristics/features to be rated (rows):** Skills, knowledge, and/or behavior to be demonstrated;
3. **Levels of mastery/scale (columns):** Labels used to describe the levels of mastery; these should be direct and clear;
4. **A description of each characteristic at each level of mastery/scale (cells).**

Using rubrics offers many significant advantages. They provide students with a clear understanding of the expectations by which their work will be assessed and evaluated. The constructive feedback rubrics help students improve the quality of their performance and become self-directed and self-assessing learners (Huba & Freed, 2000). For teachers, rubrics help determine in advance the criteria necessary for specific desired learning outcomes. This allows them to concentrate on what is necessary to ensure that students are engaged in attaining these outcomes.

In this study, student participants were asked to design an educational product from the materials available on the WebQuest site as a part of their course assessment requirement. The design and description are discussed in the Instrumentation section.

RESEARCH QUESTIONS AND HYPOTHESES

The current study focused on two research questions:

- RQ1:** Do alternative assessment approaches have an effect on the quality of WebQuest design?
RQ2: Are there differences in the scores given by the teacher and the students assessing the quality of WebQuest designs?

To answer these two questions, two alternative hypotheses were formulated:

- There are statistically significant differences at the 0.05 level between the means of assessment scores in terms of assessment routes (route 1, route 2) of the quality of WebQuest design;
- There are statistically significant differences at the 0.05 level of the means of assessment scores between teacher and student in both the self- and peer- assessment approaches of the quality of WebQuest design.

METHODOLOGY

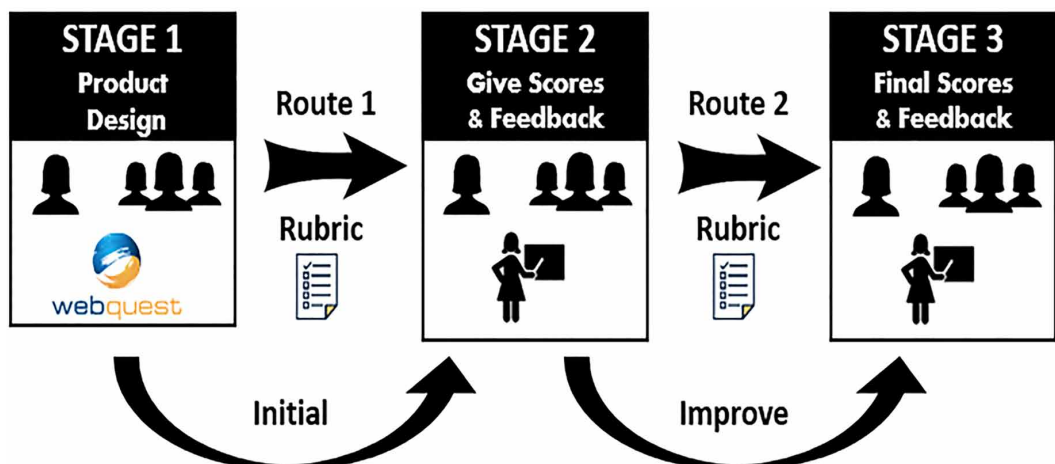
Participants

This study used a probability (random) sampling approach, which is a widely-used strategy in quantitative studies (Cohen, Manion, & Morrison, 2002). Participants were chosen from graduates enrolled in a course entitled “Computer Applications in Education,” which is a compulsory course in the ‘High Diploma in Education Program’ offered by the College of Education at King Faisal University (KFU). Participants had finished their bachelor’s degrees and enrolled in this program, which is considered to be a teacher preparation program for graduates seeking to qualify as future teachers. Participants were females from different age groups ranging from 22-35 years old. They came from different majors and specialties, having graduated from different colleges at KFU, such as the College of Education, the College of Science, the College of Art, and the College of Computer Sciences. There were seven classes in this course, of which two were randomly chosen to take part in this study. Class A and Class B consisted of 24 students each, with a total of 48 students participating. All participants were of the same academic level. The students were asked to participate in this study as partial fulfillment of their “Computer Applications in Education” course.

Intervention

The intervention of this study was implemented in three stages, as illustrated in Figure 1. The two routes [sequences] in the figure refer to assessment before feedback and after feedback.

Figure 1. The steps in implementing alternative assessments. Source: (Al-Abdullatif, 2020).



Stage One: The students were oriented to the concept of WebQuests, learned about the six principal components of the program, and were shown how to design them using Google Sites. Then the students were divided into two groups: one used the self-assessment approach and the other

used the peer-assessment approach. These groups were asked to design and produce WebQuests and were given an assessment rubric. How the items in the rubric should be used in the first assessment stage of their products was explained to the students.

Stage Two: The second stage represented the first route in this process. The instructor gave the students an assessment rubric link via the BLMS and asked the two groups (self and peer) to evaluate the WebQuests. All participants were asked to upload their WebQuest links on the discussion board, via the BLMS. The 24 students in the self-assessment group each evaluated her own WebQuest. The peer-assessment group was composed of 24 students. Since it would have been difficult for each student to evaluate 23 WebQuests, the students were divided into four groups with six students in each. Each student was asked to evaluate the five WebQuests of the other members of her group and post their feedback upon each WebQuest via the discussion board.

The instructor evaluated the WebQuests of the self-assessment group and the peer-assessment group at the same time using the same assessment rubric, and provided feedback to each group. By comparing the instructor's scores and the scores given by the students, the instructor discussed with them the differences between their assessment of their products and his assessment, and how to improve the quality of their WebQuest designs. In this stage, the students experienced using the alternative assessment tools of self and peer-assessment for the first time.

Stage Three: The third stage represented the second route in this approach. Having received feedback in the second stage, the students modified, improved, and developed their WebQuest designs, following the same steps described in stage two: self-assessment or peer-assessment, instructor's assessment, and feedback. In this stage, the students' final scores were calculated by the instructor for both the self and peer-assessment groups and were compared with the student assessment scores.

Instrumentation

In this study, an analytic assessment rubric was used to assess student performance in product design, specifically a WebQuest project. To assess the quality of the student WebQuests, a rubric was designed based on a review of the related literature (Dodge, 2001; eMINTS, 2006; Unal, Bodur, & Unal, 2012; March, 2018).

This rubric consisted of four parts within a matrix:

1. **The WebQuest task description:** This part, describing the criteria, was located in the first column of the matrix and included the following seven items:
 - a. Overall design, which included three elements: the general appearance of the project, spelling and grammatical mistakes, and internal navigation;
 - b. Introduction, which included two elements: stimulation and cognitive efficiency;
 - c. Tasks, which involved one element: the clarity of the task;
 - d. Operating procedures, which included three elements: the clarity of the process, the cognitive level, and cooperation;
 - e. Resources, which included two elements: the amount of resources and their quality;
 - f. Evaluation, which included one element: clarity of the evaluation criterion;
 - g. Conclusion, which included one element: relation the final message of the WebQuest to the student knowledge and to experiences they might encounter in other educational situations;
2. **The characteristics to be rated:** This part was located in the rows above each item of the WebQuest tasks that described the skills required to satisfy each of the criteria;
3. **Scale:** This part was represented in the three columns of the WebQuest matrix that described the level of mastery of each item. A rating scale of three levels (Low, Intermediate, and Excellent) was based on scores from 0 points–4 points. This score rating varied from item to item;

4. **A description of each characteristic at each level of the scale:** This part was represented in all the cells of the WebQuest matrix.

The validity of the assessment rubric was achieved by seeking the opinion of a group of experts in instructional technology and educational psychology. They were consulted about how well the seven domains and score levels of the assessment rubric reflected the subject of the instrument, whether the domains were described accurately, and whether they were appropriate. The experts generally approved of the instrument; however, they recommended that some expressions be reformulated and others deleted. After the author had followed both suggestions, the assessment rubric was in its final form and ready to be applied to the pilot study participants.

According to Moskal (2000), interrater reliability, which refers to the consistency of scores assigned by multiple evaluators or raters, is relevant to the scoring of rubrics. The reliability of the assessment rubric was measured for test-retest reliability using the Spearman-Brown coefficient on a pilot sample of 20 students representative of the study's population. Groups of two students participated in designing one WebQuest. A total of 10 WebQuests were designed and were then assessed by the two teachers, the raters in this case. The Spearman-Brown coefficient was calculated to determine the correlation between two administrations of the same assessment rubric by two different teachers who were evaluating the same WebQuest designs. The internal consistency of the assessment rubric was 0.89, which, according to Field (2013), indicates a high reliability.

Data Collection and Procedures

The study was implemented during the second semester of the 2017-18 school year, from January to March. The teacher divided the participants into two groups during the study intervention. An introductory lecture was held to provide the students with a concise introduction to the WebQuest concept, objectives, components, and how a WebQuest is designed on Google Sites. The assessment rubric and what they were expected to achieve by the end of the project were also explained to them.

After the participants had experienced the three stages of the intervention, the assessment scores from both the students and the teacher were collected and documented. In the case of the self-assessment approach, student scores were compared with the teacher's scores. In the case of the peer-assessment approach, the assessment score of each student was calculated by the mean scores of the five students in the same group. That is, each student had an individual score that was compared with the teacher's scores.

SPSS v.23 was used to conduct the statistical analysis. Descriptive statistical analyses of the means, standard deviation, and standard error of the assessment scores were performed for the eight groups of the intervention (teacher in self-assessment approach in route 1 and route 2; students in self-assessment approach in route 1 and route 2; teacher in peer-assessment approach in route 1 and route 2; and students in peer-assessment approach in route 1 and route 2). The descriptive analysis aimed to provide general insights into the participants' assessment scores.

To assess the impact of the alternative assessment approaches on the quality of product design, an independent samples t-test was performed that measured the statistical differences between the means of the eight groups in both route 1 and route 2.

A one-way ANOVA was performed in order to allocate the statistical differences between teacher and student in both the self-and peer-assessment approaches. Scheffe's coefficient was used to detect where the significance lay and in which favor. For the t-test and the one-way ANOVA, the 0.05 alpha level was used as the criterion for statistical significance. The results are discussed in the following section.

FINDINGS AND DISCUSSION

Findings are discussed in terms of the two research questions of this study:

RQ1: Did alternative assessment approaches affect the quality of WebQuest design?

To answer this question, an alternative hypothesis was formulated: There are statistically significant differences at the 0.05 level between the means of assessment scores of the quality of WebQuest design in terms of assessment routes (route 1, route 2). Therefore, to test the validity of this hypothesis, an independent sample t-test was performed.

Table 1 shows the mean differences of assessment scores between the two routes (route 1, route 2) of the intervention for all four group participants (teacher in the self-assessment approach, student

Table 1. Independent samples t-test of assessment scores for both teacher and students between route 1 and route 2 (N = 48)

Groups	Route 1		Route 2		T-Test for Equality of Means		
	M	SE	M	SE	t	df	Sig.
Teacher-Self	16.5	0.56	25.5	0.23	14.805	46	0.000
Student-Self	23.0	0.32	26.7	0.22	9.365	46	0.000
Teacher-Peer	19.0	0.76	24.3	0.22	6.570	46	0.000
Student-Peer	18.5	0.47	23.0	0.39	7.296	46	0.000

Source: (Al-Abdullatif, 2020)

in the self-assessment approach, teacher in the peer-assessment approach, and student in the peer-assessment approach).

The data indicates that there are statistically significant differences for each group between route 1 and route 2 in favor of the second route of the intervention. In the first group, the teacher scored the quality of students' WebQuest designs in the self-assessment approach environment; the data shows significant differences in means between the teacher's scores in route 1 ($m = 16.5$) and route 2 ($m = 25.5$), $t(46) = 14.805$, $p < 0.05$. In the second group, the students scored themselves on the quality of their WebQuest designs in the self-assessment approach environment; the data shows significant differences in means between the student scores in route 1 ($m = 23.0$) and route 2 ($m = 26.7$), $t(46) = 9.365$, $p < 0.05$. In the third group, the teacher scored the students' quality of WebQuest designs in the peer-assessment approach environment; the data shows significant differences in means between the teacher's scores in route 1 ($m = 19.0$) and route 2 ($m = 24.3$), $t(46) = 6.570$, $p < 0.05$. In the fourth group, the students scored their peers on the quality of their WebQuest designs in the peer-assessment approach environment; the data shows significant differences in means between the student scores in route 1 ($m = 18.5$) and route 2 ($m = 23.0$), $t(46) = 7.296$, $p < 0.05$.

Therefore, the data generally indicates the significant impact of alternative assessment approaches on student performance, an impact represented in the significant increase of the assessment scores on the quality of their WebQuest designs. This is due to student awareness of the assessment criteria rubric provided prior to feedback and to the fact that, between the first route and second route of the assessment, relative and constructive feedback was offered by both students and their teacher about how they could improve the quality of their products. This result supports the literature, which has reported that alternative assessment approaches have a positive impact on learning and achievement (Leach, 2012; Liu, & Lee, 2013; Ashraf & Mahdinezhad, 2015; Zarei & Usefli, 2015; Irajii et al., 2016; McKeivitt, 2016).

RQ2: Are there any differences between the teacher’s and students’ assessment scores of the quality of WebQuest design?

To answer this question, an alternative hypothesis was formulated: There are statistically significant differences between the means of teacher’s and students’ assessment scores of the quality of WebQuest design at the 0.05 level of in both the self-and peer-assessment approaches. To test the validity of this hypothesis, a one-way ANOVA was performed.

Table 2 provides the means, standard deviations, and standard errors of the eight independent groups. The first four groups are the assessment scores of the teacher in both routes of the intervention

Table 2. Descriptive data of the four independent groups (N = 24)

Groups	M	SD	SE
Teacher-Self 1	16.5	2.75	0.56
Teacher-Self 2	25.5	1.14	0.23
Teacher-Peer 1	19.0	3.75	0.77
Teacher-Peer 2	24.3	1.11	0.23
Student-Self 1	23.0	1.61	0.33
Student-Self 2	26.7	1.11	0.23
Student-Peer 1	18.5	2.34	0.48
Student-Peer 2	23.0	1.91	0.39

Source: (Al-Abdullatif, 2020)

(route 1, route 2) and in the two assessment approaches (self, peer). The second four groups represent the assessment scores of the students in both routes of the intervention (route 1, route 2) and in the two assessment approaches (self, peer).

To determine the statistical differences in means among the eight groups of the intervention, a one-way ANOVA statistical analysis was performed. Its results, presented in Table 3, reveal that there

Table 3. Differences of assessment scores between groups using one-way ANOVA (N = 48)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2240.250	7	320.036	68.873	.000
Within Groups	855.000	184	4.647		
Total	3095.250	191			

Source: (Al-Abdullatif, 2020)

is a statistical significance among the eight groups. Given that $p < 0.05$, the alternative hypothesis was accepted: There are statistically significant differences at the 0.05 level of the means of assessment scores between teacher and students in both the self and peer-assessment approaches of the quality of WebQuest design. $F(7,184) = 68.873, p < 0.05$.

To detect where the significance lies and in which favor, that is, among which groups of the eight a significant difference in teacher and student assessment scores can be seen, the Scheffe coefficient was used (Table 4).

Table 4. Multiple comparisons between groups using the Scheffe coefficient (N = 48)

Groups	Teacher-Self 1	Teacher-Self 2	Teacher-Peer 1	Teacher-Peer 2
Student-Self 1	0.000*	0.000*	0.000*	0.775
Student-Self 2	0.000*	0.775	0.000*	0.000*
Student-Peer 1	0.178	0.000*	0.999	0.000*
Student-Peer 2	0.000*	0.775	0.000*	0.775

*The mean difference is significant at the 0.05 level
 Source: (Al-Abdullatif, 2020)

As shown in Table 4, there is a significant difference in the assessment scores of students assessing themselves (self-assessment) in route 1 of the invention ($m = 23$) and the assessment scores of the teacher assessing the same students in the self-assessment approach ($m = 16.5$) in favor of the group of higher mean. This result indicates that students engaging in the self-assessment approach in route 1—where they experienced the self-assessment intervention for the first time—was subjective and highly biased. This result is supported by the study of Ashraf and Mahdinezhad (2015, p. 111) that found “self-assessment has a risk of being perceived as a process of presenting inflated grades and being unreliable.” However, most of the studies in the related literature contradict this result, instead emphasizing the objectivity and reliability of the self-assessment method (Xiao & Lucking, 2008; Chen & Tsai, 2009; De Wever et al., 2011; Pöldoja et al., 2014; Uto & Ueno, 2016). Moreover, several studies have indicated that self-assessment is more effective than teacher-assessment in improving student performance (Esfandiari & Myford, 2013; Zarei & Usefli, 2015). The high degree of bias in this result is due to the fact that the self-assessment experience was a new approach for the participants. The traditional context of educational assessment depends on assessments by teachers only, which led to the students’ inability to assess themselves objectively and to make credible judgments.

There is no significant difference in the scores of students assessing themselves (self-assessment) in route 2 of the intervention ($m = 26.7$) and the assessment scores of the teacher assessing the same students in the self-assessment approach ($m = 25.5$) in the same route. The convergence of student and teacher results in the route 2 indicates the increase of students’ objectivity when assessing themselves for the second time; it suggests that they had become more neutral and thoughtful in assessing themselves according to the criteria given to them. This effect is a product of the self-assessment approach, which improves student awareness of being objective when assessing themselves, and of the feedback and guidance received from their teacher. Together, these factors improve the quality of the product, as shown in the first hypothesis (Leach, 2012; Liu & Lee, 2013; Ashraf & Mahdinezhad, 2015; Zarei & Usefli, 2015; Irajii et al., 2016; McKeivitt, 2016).

There was no significant difference in the assessment scores of students assessing their peers (peer-assessment) in route 1 of the intervention ($m = 18.5$) or the assessment scores of the teacher assessing the same students in the peer-assessment approach ($m = 19.0$) in the same route. In addition, there was no significant difference in the assessment scores of students assessing their peers (peer-assessment) in route 2 of the intervention ($m = 23.0$) or the assessment scores of the teacher assessing the same students in the peer-assessment approach ($m = 24.3$) in the same route. This suggests that the peer-assessment mode was characterized by greater objectivity and impartiality than the self-assessment mode throughout the period of the intervention. This finding is supported by similar studies in the related literature (Khonbi & Sadeghi, 2013; Zarei & Mahdavi, 2014; Ashraf & Mahdinezhad, 2015).

The second route of the application of alternative assessment methods occurred during the stage in which the students were more experienced in self and peer-assessment than they had been during the first stage, when they were new to such experiences. The result generally proves that both the self-assessment and peer-assessment approaches are effective when assessing the quality of the educational products of students.

CONCLUSION AND IMPLICATIONS

The purpose of this study was to investigate the impact of web-based alternative assessment approaches, both self-assessment and peer-assessment, on improving the quality of student-designed educational products, specifically WebQuest projects as a part of a course requirement. By assessing the quality of educational products the students designed, the reliability and objectivity of the two assessment approaches could be determined.

The findings provide guidelines for teachers on the design and use of alternative web-based assessment approaches which can improve the quality of educational product design and better support its assessment. The study focused on auditing the credibility and objectivity of two assessment approaches by comparing the scores of student assessments with the teacher's assessment of product quality. The results showed that the method of peer-assessment is characterized by high credibility and objectivity regardless of the students' previous experience in assessing their peers. Self-assessment, however, possesses a high degree of bias and subjectivity. For this reason, the study emphasized the determining factor of student experience with this approach, and the vital role this experience plays in improving their performance.

The current study recommends the utilization of alternative web-based assessment approaches, as they have a significant impact on the enhancement of student learning skills and performance. These approaches also reinforce student active integration into the evaluation process more effectively than traditional teacher-based evaluation. The results of this study imply that, in web-based learning environments, training can increase the credibility and objectivity of the self-assessment method. Before the actual assessment, students should practice self-assessment techniques and be trained on how to assess themselves using the provided assessment criteria. In addition, when web-based self and peer-assessment are included in determining the quality of an educational product's design, peer-assessment should be given more weight than self-assessment due to its objectivity and consistency with teacher-assessment. However, self-assessment remains a valuable learning activity, especially when paired with peer-assessment activities, as the multi-dimensional assessment processes are necessary to increase assessment validity and reliability (Chang et al., 2012). The study recommends the need to assess the impact of other alternative web-based assessment methods and techniques, particularly in terms of improving student performance when designing educational products. The validity and reliability of these methods could also be studied in future research. Finally, it will also be important to evaluate the impact of web-based self and peer-assessment on other learning outcomes, such as self-regulated learning skills, self-efficacy, and learning styles.

ACKNOWLEDGMENT

This research was supported by the Deanship of Scientific Research at King Faisal University [grant number 186049].

REFERENCES

- Al-Abdullatif, A. M. (2012). *An investigation into the perceptions of university students and instructors on the effectiveness of online education in a Saudi tertiary environment*. Unpublished Doctoral dissertation, Griffith University.
- Andrade, H. G., & Boulay, B. A. (2003). Role of rubric-referenced self-assessment in learning to write. *The Journal of Educational Research*, 97(1), 21–30. doi:10.1080/00220670309596625
- Andrade, H. G., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, 32(2), 159–181. doi:10.1080/02602930600801928
- Ashraf, H., & Mahdinezhad, M. (2015). The role of peer-assessment versus self-assessment in promoting autonomy in language use: A Case of EFL Learners. *Iranian Journal of Language Testing*, 5(2), 110–120.
- Atan, H., Rahman, Z. A., & Idrus, R. M. (2004). Characteristics of the web-based learning environment in distance education: Students' perceptions of their learning needs. *Educational Media International*, 41(2), 103–110. doi:10.1080/09523980410001678557
- Biggs, J., & Moore, P. (1993). *The process of learning* (3rd ed.). New York: Prentice Hall.
- Birjandi, P., & Bolghari, M. S. (2015). The relationship between the accuracy of self-and peer-assessment of Iranian intermediate EFL learners and their learning styles. *Theory and Practice in Language Studies*, 5(5), 996–1006. doi:10.17507/tpls.0505.15
- Boud, D. (2013). *Enhancing learning through self-assessment*. New York: Routledge. doi:10.4324/9781315041520
- Brophy, S. T. (2012). Writing effective rubrics: Institutional assessment. Retrieved from http://web.cse.ohio-state.edu/~soundarajan.1/abet/writing_effective_rubrics_guide_v2.pdf
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245–281. doi:10.3102/00346543065003245
- Butler, Y., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing*, 27(1), 5–31. doi:10.1177/0265532209346370
- Chang, C. C., Tseng, K. H., & Lou, S. J. (2012). A comparative analysis of the consistency and difference among teacher-assessment, student self-assessment and peer-assessment in a web-based portfolio assessment environment for high school students. *Computers & Education*, 58(1), 303–320. doi:10.1016/j.compedu.2011.08.005
- Chen, Y. C., & Tsai, C. C. (2009). An educational research course facilitated by online peer-assessment. *Innovations in Education and Teaching International*, 46(1), 105–117. doi:10.1080/14703290802646297
- Cheng, K. H., Liang, J. C., & Tsai, C. C. (2015). Examining the role of feedback messages in undergraduate students' writing performance during an online peer-assessment activity. *The Internet and Higher Education*, 25, 78–84. doi:10.1016/j.iheduc.2015.02.001
- Cohen, L., Manion, L., & Morrison, K. (2002). *Research methods in education*. London: Routledge. doi:10.4324/9780203224342
- De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2011). Assessing collaboration in a wiki: The reliability of university students' peer-assessment. *The Internet and Higher Education*, 14(4), 201–206. doi:10.1016/j.iheduc.2011.07.003
- Dodge, B. (2001). *A rubric for evaluating WebQuests*. Retrieved from <http://webquest.org/sdsu/webquestrubric.html>
- Domínguez, C., Jaime, A., Sánchez, A., Blanco, J. M., & Heras, J. (2016). A comparative analysis of the consistency and difference among online self-, peer-, external-and instructor-assessments: The competitive effect. *Computers in Human Behavior*, 60, 112–120. doi:10.1016/j.chb.2016.02.061
- Elliott, R. (2008). Assessment 2.0. *International Journal of Emerging Technologies in Learning*, 3, 66–70.
- eMINTS. (2006). *Rubric/scoring guide for WebQuests*. Retrieved from <http://emints.org/webquest/rubric.shtml>

- Esfandiari, R., & Myford, C. M. (2013). Severity differences among self-assessors, peer-assessors, and teacher assessors rating EFL essays. *Assessing Writing, 18*(2), 111–131. doi:10.1016/j.asw.2012.12.002
- Falchikov, N. (2001). *Learning together: Peer tutoring in higher education*. London: Routledge.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Los Angeles: SAGE.
- Gibbs, G. (2010). *Using assessment to support student learning*. Leeds: Leeds Met Press.
- Hsu, Y. C., Ching, Y. H., & Grabowski, B. L. (2014). Web 2.0 applications and practices for learning through collaboration. In *Handbook of research on educational communications and technology* (pp. 747–758). New York, NY: Springer. doi:10.1007/978-1-4614-3185-5_60
- Huba, M., & Freed, J. (2000). *Learner-Centered Assessment on College Campuses*. Boston: Allyn & Bacon.
- Iraji, H. R., Enayat, M. J., & Momeni, M. (2016). The effects of self-and peer-assessment on Iranian EFL learners' argumentative writing performance. *Theory and Practice in Language Studies, 6*(4), 716. doi:10.17507/tpls.0604.08
- Kao, G. Y. M. (2013). Enhancing the quality of peer review by reducing student “free riding”: Peer-assessment with positive interdependence. *British Journal of Educational Technology, 44*(1), 112–124. doi:10.1111/j.1467-8535.2011.01278.x
- Khonbi, Z. A., & Sadeghi, K. (2013). The effect of assessment type (self vs. peer) on Iranian university EFL students' course achievement. *Procedia: Social and Behavioral Sciences, 70*, 1552–1564. doi:10.1016/j.sbspro.2013.01.223
- Kilimci, S. (2010). Integration of the internet into a language curriculum in a multicultural society. *Turkish Online Journal of Educational Technology-TOJET, 9*(1), 107–113.
- Leach, L. (2012). Optimal self-assessment: Same tensions and dilemmas. *Assessment & Evaluation in Higher Education, 37*(2), 137–147. doi:10.1080/02602938.2010.515013
- Liu, E. Z. F., & Lee, C. Y. (2013). Using peer feedback to improve learning via online peer-assessment. *Turkish Online Journal of Educational Technology-TOJET, 12*(1), 187–199.
- Liu, E. Z. F., Shih, R. C., & Tsai, Y. L. (2011). Hyperlink network analysis of the educational blog. *British Journal of Educational Technology, 42*(2), E25–E29. doi:10.1111/j.1467-8535.2010.01150.x
- Liu, N. F., & Carless, D. (2006). Peer feedback: The learning element of peer-assessment. *Teaching in Higher Education, 11*(3), 279–290. doi:10.1080/13562510600680582
- McKevitt, C. T. (2016). Engaging students with self-assessment and tutor feedback to improve performance and support assessment capacity. *Journal of University Teaching & Learning Practice, 13*(1), 2.
- Moskal, B. M. (2000). Scoring Rubrics: What, when and how? *Practical Assessment, Research Evaluation, 7*(3), 1–5.
- Põldoja, H., Väljataga, T., Laanpere, M., & Tammets, K. (2014). Web-based self-and peer-assessment of teachers' digital competencies. *World Wide Web (Bussum), 17*(2), 255–269. doi:10.1007/s11280-012-0176-2
- Price, M., Carroll, J., O'Donovan, B., & Rust, C. (2011). If I was going there I wouldn't start from here: A critical commentary on current assessment practice. *Assessment & Evaluation in Higher Education, 36*(4), 479–492. doi:10.1080/02602930903512883
- Sitthiworachart, J., & Joy, M. (2003). Web-based peer-assessment in learning computer programming. In *Proceedings of the 3rd IEEE International Conference on Advanced Learning Technologies* (pp. 180–184). IEEE Press. doi:10.1109/ICALT.2003.1215052
- Spiller, D. (2012). *Assessment matters: Self-assessment and peer-assessment*. Hamilton, New Zealand: The University of Waikato.
- Sung, Y. T., Chang, K. E., Chiou, S. K., & Hou, H. T. (2005). The design and application of a web-based self-and peer-assessment system. *Computers & Education, 45*(2), 187–202. doi:10.1016/j.compedu.2004.07.002

Taş, E., Cetinkaya, M., Karakaya, Ç., & Apaydin, Z. (2013). An investigation on web designed alternative measurement and assessment approach. *Education & Science/Eğitim ve Bilim*, 38(167), 196-210.

Tom March. (2018). *Assessing best WebQuests*. Retrieved from <http://tommarsh.com/strategies/webquests/assessing-best-webquests/>

Unal, Z., Bodur, Y., & Unal, A. (2012). A standardized rubric for evaluating WebQuest design: Reliability analysis of ZUNAL WebQuest design rubric. *Journal of Information Technology Education: Research*, 11, 169–183. doi:10.28945/1688

University of Hawai'i. (2017). *Assessment how-to*. Retrieved from <https://manoa.hawaii.edu/assessment/howto/rubrics.htm>

Uto, M., & Ueno, M. (2016). Item response theory for peer-assessment. *IEEE Transactions on Learning Technologies*, 9(2), 157–170. doi:10.1109/TLT.2015.2476806

Vickerman, P. (2009). Student perspectives on formative peer-assessment: An attempt to deepen learning. *Assessment & Evaluation in Higher Education*, 34(2), 221–230. doi:10.1080/02602930801955986

Xiao, Y., & Lucking, R. (2008). The impact of two types of peer-assessment on students' performance and satisfaction within a wiki environment. *The Internet and Higher Education*, 11(3-4), 186–193. doi:10.1016/j.iheduc.2008.06.005

Zarei, A. A., & Sayar Mahdavi, A. (2014). The effect of peer and teacher-assessment on EFL learners' grammatical and lexical writing accuracy. *Journal of Social Issues and Humanities*, 2(9), 92–97.

Zarei, A. A., & Usefli, Z. (2015). The Effect of Assessment Type on EFL Learners' Goal-Orientation. *Journal of Language. Linguistics and Literature*, 1(4), 112–119.

Ahlam Mohammed Al-Abdullatif is an assistant professor in Educational Technology Department at King Faisal University in Saudi Arabia. Al-Abdullatif's research interest is in the area of eLearning particularly on how Information and Communication Technology (ICT) enhances teaching and learning practices. She is an active professor and researcher and published a number of research articles with well-known publishers such as Taylor and Francis and Springer. Al-Abdullatif won the 2017/2018 Faculty Excellence Award of KFU and the 2017/2018 E-learning Excellence Award of the National Center of E-learning in Saudi Arabia.