

# A Review of the Literature on Online (e) Assessment

*Kevin Mawhinney, Technology Education Department Head, Cobequid Educational Centre, Truro, Nova Scotia, mawhinneyk@ccrsb.ca*

## **Introduction**

Gaytan and McEwen (2009) defined online assessment as a system for evaluating students' academic achievement in an electronic environment. Walker, Topping and Rodrigues (2008) noted that the purpose of online assessment is to monitor student understanding, improve academic programs, and enhance student learning. The authors observed that e-assessment is an important theme for researchers because the use of online assessment is increasing rapidly yet research surrounding its use is limited. According to Dermo (2009), e-assessment is widely recognized as a key issue in improving the quality of students' e-learning. Gaytan and McEwen argued that it is incumbent upon teachers to become knowledgeable about online assessment as current educational reform movements place increased demands for accountability, improvement, and achievement on online practitioners.

This paper presents a review of ten studies on online assessment. The review includes analysis and critique of the findings of the 10 studies as well as implications, and limitations of the analysis. The paper begins with an overview of how the studies were selected and reviewed.

## **Methods**

Studies were identified by conducting keyword searches in academic indexes using both Education Resources Information Center (ERIC) and Google Scholar. Because sources were selected from educational technology journals and books, all included an electronic medium. The analysis only included sources with the words "online assessment", "e assessment", or "web assessment" as part of the title. To figure as part of the analysis, the journal sources had to include research participants. This means that meta-analyses, book reviews etc were excluded. A number of sources were excluded because the primary focus was not on student assessment in online environments (e.g.,: comprehensive geriatric assessment online). Due to the emerging nature of e-assessment, more recent studies were favored for inclusion over older studies. The studies selected ranged from 2004 to 2011 and were taken from the peer-reviewed journals listed in Table 1.

Table 1

*Journals Included in the Literature Review*

Number of studies	Journal Name
1	Educational Technology & Society
1	Innovations in Education and Teaching International
1	Computers & Education
1	British Journal of Educational Technology
2	The American Journal of Distance Education
1	Canadian Journal of Learning and Technology
1	Journal of Distance Education
1	Learning, Media and Technology
1	Journal of Instructional Technology and Distance Learning

The studies were conducted in the USA, United Kingdom, Turkey, Australia, Algeria and Taiwan. Seven of the studies used undergraduate student as participants, three used graduate students, and two used University faculty members. There was a broad range in the number of participants involved in the research. Two studies used less than 20 participants, five studies used between 30 and 52 participants, and four studies used in excess of 200 participants. Four studies used a quantitative approach, three used qualitative, and three used mixed methods. Table 2 describes the studies.

Table 2

*Descriptive Characteristics of the Studies*

Authors	Country	Sample size	Participants	Method
Bouzidi & Jaillet (2009)	Algeria	242	Undergraduate students	Quantitative
Chen & Tsai (2009)	Taiwan	52	Graduate Students	Quantitative
Costa, Mullan, Kothe, & Butow (2010)	Australia	43	Undergraduate students, faculty members, and graduate students	Quantitative
Dermo (2009)	UK	30	Undergraduate students	Qualitative

Gaytan & McEwen (2009)	US	2028	Undergraduate students and faculty members	Quantitative
Li, Steckelberg, & Srinivasan (2008)	USA	38	Undergraduate students	Mixed method
Ozden, Erturk, & Sanli (2004)	Turkey	46	Undergraduate students	Qualitative
Walker, Topping, & Rodrigues (2008)	UK	15	Undergraduate students	Qualitative
Yates & Beaudrie (2009)	USA	850	Undergraduate Students	Mixed method
Yilmaz (2010)	Turkey	14	Graduate students	Mixed method

The studies were analyzed to identify similarities and differences, patterns and themes. The analytical review of the literature was consistent with a qualitative meta-synthesis method that uses findings from existing studies as data to build new understanding (Zimmer, 2004). Zimmer added that the method allows for, “the identification of consensus, hypothesis development, and investigation of contradictions in patterns of experience across studies” (p. 312).

## **Findings**

Four themes emerged from the findings of the 10 studies as follows: perceptions of e-assessment; validity and reliability of e-assessment; supporting students in e-assessment; and the benefits e-assessment. Each theme is reported on separately below.

### ***Perceptions of e-assessment***

Seven of the studies inquired into student perceptions and attitudes towards online assessment. Dermo (2009) found that students perceived e-assessment positively in most categories including reliability, security, validity, accessibility, and as adding value to their learning. Demo concluded that the perception of e-assessment was consistent regardless of the participants’ age and gender. Participants perceived as unfair the assignment of random questions from an item bank as they did not feel that the questions were of equal difficulty.

Dermo (2009) and Walker, Topping, and Rodrigues (2008) both noted that students perceived formative e-assessment as less stressful than alternative formats while demanding a similar cognitive level. Costa, Mullan, Kothe and Butow (2010) found that, overall, 67% of participants perceived some benefit from using online computer-based assessment with 58% of participants agreeing that they enjoyed using the method. Ninety-two percent of post-graduate participants completed at least one online quiz suggesting that they perceived formative online assessment as beneficial.

Ozden, Erturk, and Sanli (2004) found that the majority of students perceived formative online assessment as fair, as capable of providing effective feedback, and as a stimulus that caused them to reflect on their learning. Eighty percent of participants perceived online assessment as more contemporary and systematic than conventional assessment methods. Participants were dissatisfied that it was not possible to edit prior answers since they were hidden on completion to prevent cheating.

Walker, Topping, and Rodrigues (2008) studied student perceptions of formative e-assessment. Findings revealed that participants perceived formative e-assessment as a useful tool that allowed them to assess personal progress and identify learning needs. A minority of participants perceived e-assessment as a more mechanical tool to achieve outcomes and improve grades. Participants felt that allowing unlimited assessment attempts was ineffective and promoted shallow learning. Unprompted questions such as fill in the blanks were perceived as unjust as they required too great a degree of perception. The majority of participants accepted e-assessment as a useful learning aid, saw it as less stressful than alternative forms of assessment, and felt that it required the same cognitive level as conventional tests. The presence of feedback during e-assessment was viewed positively as it reinforced learning and provided deeper insight.

Chen and Tsai (2009) and Li, Steckelberg, and Srinivasan (2008) studied the use of online peer assessment in higher education. Chen and Tsai found that students' attitudes towards online peer assessment were generally positive prior to, and following peer assessment. Although the study did not show an increase in favorable attitudes following the treatment, participants' positive attitudes were statistically maintained. Participants in the Li, Steckelberg, and Srinivasan study also expressed positive perceptions of online peer assessment. The majority of students felt that peer assessment was a worthwhile activity, that they benefited from peers' comments, from marking the work of their peers, and that their projects improved because of peer assessment. Some participants in the study felt that anonymous peer assessment encouraged overly critical comments. The authors found that students had a positive perception of online peer assessment when peer pressure was controlled, feedback was accessed in a timely manner, and when sufficient training and a marking key were provided.

Yilmaz (2010) studied student perceptions of peer assessment as a model for evaluating online group work. The authors found that peer assessment was positively perceived for its ability to foster collaboration, promote effective feedback, and illustrate the diversity that exists within the class. Gaytan and McEwen (2007) administered

surveys to faculty and students to measure which assessment strategies were perceived as most effective. Faculty members identified effective assessment strategies as rubrics, peer evaluation, threaded discussions, online chats and timed quizzes, portfolios, and self-assessment. Students perceived the most effective e-assessment tools as self assessments and practice tests, threaded discussions, weekly assignments with immediate feedback, the use of rubrics, and portfolios. Students indicated that e-assessment could be improved by providing meaningful and timely feedback, utilizing a well-designed rubric, and employing a variety of assessment techniques.

### ***Validity and reliability of online assessment***

Four of the studies examined the validity of online assessment. Bouzidi and Jaillet (2009) noted that, individually, students are not considered as reliable assessors. The authors conducted a study to determine if several students would collectively constitute a valid assessor. They found a high correlation between teacher assessment and peer assessment. The evidence confirmed that peer assessment is valid when it is applied to exact science fields and when the assessment is marked by more than four peers. The study also revealed that the validity of peer assessment could be improved by grading students' participation in peer assessment and by combining online peer assessment with self assessment.

Chen and Tsai (2009) studied the validity of online peer assessment. Their findings were consistent with those of Bouzidi and Jaillet who noted that peer grades were highly correlated with instructors' scores. Yates and Beaudrie (2009) studied the impact that unsupervised online assessment has on student grades to determine whether the results differ from those acquired by online students taking exams in an in-person, proctored environment. The authors found no significant difference in grades between the two conditions suggesting that the online results are as valid as those conducted in person. Yilmaz (2010) found that, to eliminate problems with reliability and validity, students must be provided with detailed rubrics, training, and support mechanisms so that they know what is expected of them and are equipped with the tools required to fulfill their responsibility.

### ***Supporting students in e-assessment***

Six of the studies found that students require ongoing support to facilitate their participation in online assessment. Bouzidi and Jaillet (2009), Chen and Tsai (2009), and Walker, Topping, and Rodrigues (2008) found that quality marking criteria are critical for the success of online peer and self-assessment. Since students have little experience in the role of assessor, teachers must provide support in the form of clear marking instructions and detailed rubrics. Yilmaz (2010) found that some participants lacked the critical assessment skills required to assess their peers and identify their shortcomings. As a result, they focused on the mechanics of writing only and avoided

more contentious issues. Yilmaz concluded that students must be provided with assessment tutorials and instruction related to constructive criticism and timely feedback in order to prepare them as partners in e-assessment.

Li, Steckelberg, and Srinivasan (2008) found that participants preferred to remain anonymous. The authors noted that instituting anonymous peer assessment maintains confidentiality, promotes more honest feedback, and eliminates the effects of peer pressure. Participants recognized the positive influence of training on peer assessment and commented that they benefited from learning how to conduct critical assessment and implement peer feedback.

In researching the use of online formative assessment systems, Ozden, Erturk, and Sanli (2004) found that low-attaining students and those with reduced computer access might experience anxiety before adapting to online assessment systems. The authors noted that students should receive training to ensure that they have assessment tool familiarity. This prerequisite ensures that students can focus on the content of formative e-assessment without being distracted by an unfamiliar interface.

### ***The benefits of e-assessment***

Four studies found that online assessment improved the quality of student work. In researching the effect of online peer assessment, Yilmaz (2010) found that students benefited from evaluating the work of their peers as it caused them to evaluate and reflect on their own work. The online peer assessment process provided participants with an enriching experience that promoted deeper understanding and a better appreciation of their own strengths and weaknesses. Chen and Tsai (2009) analyzed assessment scores after a series of online peer assessment activities. The authors found that scores increased after each round of assessment indicating that participants benefited from both peer and expert assessment. The authors concluded that there was a positive relationship between online peer assessment and the quality of students' work.

In researching the use of online formative assessment systems, Ozden, Erturk, and Sanli (2004) found that immediate feedback and scores motivated students and contributed positively to their achievement. Walker, Topping, and Rodrigues (2008) found that the majority of students used formative assessment as a pro-active diagnostic tool to gauge their knowledge and identifying areas for further revision. Participants used analytical techniques to answer questions and indicated that formative e-assessment reinforced their learning and allowed them to test higher-level thinking skills. Participants noted that detailed feedback was required to reinforce their learning and to identify misconceptions in their understanding.

## Discussion

The seven studies on student perceptions of online assessment yielded consistent results. With few exceptions, students had positive perceptions of online formative assessment. Participants perceived online formative assessment as fair (Ozden, Erturk, & Sanli, 2004) reliable, secure, accessible, and as adding value to their learning (Dermo, 2009; Ozden, Erturk, & Sanli, 2004). Students believe that computer based assessment is contemporary, (Ozden, Erturk, & Sanli, 2004) less stressful, (Walker, Topping & Rodrigues, 2008; Dermo, 2009) enjoyable, (Costa, Mullan, Kothe and Butow, 2010) and capable of identifying their learning needs (Dermo, 2009; Walker, Topping, & Rodrigues, 2008). Perceptions of online formative assessment remained positive regardless of the age and gender of the participant.

Students also perceive online peer assessment positively (Chen & Tsai, 2009; Li, Steckelberg, & Srinivasan, 2008; Yimaz, 2010). Participants felt that online peer assessment fostered collaboration, promoted effective feedback, illustrated diversity, (Yimaz, 2010) that they benefited from peers' comments, and that their work was improved by the process (Li, Steckelberg & Srinivasan, 2008). The positive perception of peer assessment was a consistent finding across multiple studies.

Researchers discovered opposing viewpoints regarding the suitability of anonymous peer assessors. Students in the Chen and Tsai (2009) study felt that anonymous peer assessment promoted overly critical comments. Students participating in Li, Steckelberg, and Srinivasan (2008) study expressed a belief that peer anonymity reduced peer pressure, maintained confidentiality, and allowed for more constructive and honest feedback.

Despite the overall positive perceptions of online-assessment, negative perceptions of online assessment were also observed. Students took issue with the use of fill in the blanks, the process of allowing unlimited attempts, (Walker, Topping, and Rodrigues, 2008) the assigning of random questions from an item bank, (Dermo, 2009) and the inability to edit responses during online formative assessment (Ozden, Erturk, and Sanli, 2004).

Multiple authors found that students require support mechanisms to facilitate their participation in online assessment. These mechanisms include rubrics, (Bouzidi & Jaillet, 2009) clear marking criteria, (Chen & Tsai, 2009; Walker, Topping & Rodrigues, 2008) assessment tutorials, instruction related to constructive criticism, (Yilmaz, 2010) and critical assessment (Li, Steckelberg, & Srinivasan (2008). Training is required so that students are familiar with online formative assessment systems before the process of testing begins (Ozden, Erturk, & Sanli, 2004). These findings are consistent with strategies that are widely accepted as best practices for assessment in face-to-face environments.

This literature review revealed that online assessment can be a valid strategy to monitor student learning. Online unsupervised exams are as valid as those conducted in in-

person, proctored environments (Yates & Beaudrie, 2009). Peer assessment has a high correlation with teacher assessment indicating validity (Bouzidi & Jaillet, 2009; Chen & Tsai, 2009). The validity of peer assessment can be improved by grading students' participation in the assessment process, combining online peer assessment with self-assessment, by using at least four peer assessors, (Bouzidi & Jaillet, 2009) and by using rubrics and training (Yilmaz, 2010).

This literature review also identified the benefits of online assessment. For the majority of students, formative e-assessment was motivational, (Ozden, Ertuk, & Sanli, 2004) encouraged higher order thinking, reinforced learning, and served as an effective method to identify gaps in their knowledge (Walker, Topping, & Rodrigues, 2008). For the majority of students, online peer assessment promoted deeper understanding, encouraged the evaluation and reflection on their own work, and provided them with a better appreciation of their own strengths and weaknesses (Yilmaz, 2010). Online peer assessment had a positive impact on the quality of students work (Chen & Tsai, 2009). Students have positive perceptions of both formative online assessment (Chen & Tsai, 2009; Li, Steckelberg, & Srinivasan, 2008; Yilmaz, 2010) and online peer assessment (Chen & Tsai, 2009; Li, Steckelberg, & Srinivasan, 2008; Yilmaz, 2010). Online assessment can promote higher level thinking and deeper understanding, (Yilmaz, 2010) allow students to identify gaps in their understanding, (Walker, Topping, & Rodrigues, 2008) and improve the quality of their work (Chen & Tsai, 2009).

## **Limitations**

This review was limited by the small number (n=10) of studies included. In addition, the reliability of the findings of the studies reviewed may be affected by their small sample size. Li, Steckelberg, and Srinivasan (2008) noted that their research on peer assessment was based on the results of only 38 participants. Walker, Topping, and Rodrigues (2007) cautioned that only 15 participants were involved in their investigation of formative online assessment. The generalizability of the review is also limited. Yates and Beaudrie (2009) cautioned that their research into the validity of online grades involved participants from distance-education courses in a single college. The authors noted that the use of a specific population makes it difficult to generalize the findings. Since all studies used in the literature review were conducted at the undergraduate and graduate level, the findings may not be relevant to K-12 environments.

## **Implications**

To support students in their role as assessors and improve the validity of the assessment, teachers should provide sufficient training, (Li, Steckelberg, & Srinivasan) rubrics, (Bouzidi & Jaillet, 2009) scoring criteria, (Chen & Tsai, 2009; Walker, Topping & Rodrigues, 2008), and instruction (Yilmaz, 2010). As in traditional classroom, assessment, students expect fairness (Dermo, 2009) clarity, (Chen & Tsai, 2009; Walker, Topping & Rodrigues, 2008) timely feedback, and meaningful responses

(Gaytan and McEwen, 2007) so that assessment can have a positive impact on their achievement. Assessing students online appears to be as valid as assessing students in a proctored in-person environment (Yates & Beaudrie, 2009).

Students acting as online peer assessors should be provided with ongoing support to eliminate problems with validity, control peer pressure, and improve their capacity for critical assessment. Students should also be provided with adequate training, (Chen & Tsai, 2009; Li, Steckelberg, & Srinivasan, 2008) detailed rubrics and marking criteria, (Yilmaz, 2010; Chen & Tsai, 2009; Bouzidi & Jaillet, 2009; Gaytan & McEwen, 2009) exemplars, and technical assistance (Bouzidi & Jaillet, 2009). To ease anxiety associated with providing peer feedback, instruction on team dynamics and conflict resolution should be provided before project work begins (Yilmaz, 2010).

When implementing formative e-assessment systems, practitioners should ensure that students are familiar with the assessment environment prior to its use (Ozden, Erturk, & Sanli, 2004). Teachers should be trained in preparing questions for e-assessment so that they measure the intended level of knowledge (Ozden, Erturk, & Sanli, 2004) and should perform an item analysis of random test questions to ensure a similar level of difficulty (Dermo, 2009). To encourage deep thinking, teachers should provide adequate feedback and limit the number of attempts permitted (Walker, Topping & Rodrigues, 2008). To ensure that students are engaged in online formative assessment, teachers should take steps to motivate students and pilot e-assessment systems to ensure that they enable learners (Costa, Mullan, Kothe, & Butow, 2010). Instructors should adopt a variety of e-assessment tools including the evaluation of emails, discussion boards, quizzes, projects, portfolios, self-tests, and peer assessments (Gaytan & McEwen, 2009).

## References

- Bouzidi L., & Jaillet, A. (2009). Can online peer assessment be trusted? *Educational Technology & Society*, 12(4), 257-268. Retrieved from [http://www.ifets.info/journals/12\\_4/22.pdf](http://www.ifets.info/journals/12_4/22.pdf)
- Chen, Y., & Tsai, 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
- Costa, D., Mullan, B., Kothe, E., & Butow, P. (2010). A web-based formative assessment tool for Masters students: A pilot study. *Computers & Education*, 54(1), 1248-1253.
- Dermo, J. (2009). E-Assessment and the student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology*, 40(2), 203-214. doi:10.1111/j.1467-8535.2008.00915.x
- Gaytan, J., & McEwen, B. (2009). Effective online instructional and assessment strategies. *The American Journal of Distance Education*, 21(3), 117-132. Retrieved from <http://edtech.boisestate.edu/elearn/assessment.pdf>

- Li, L., Steckelberg, A., & Srinivasan, S. (2008). Utilizing peer interactions to promote learning through a web-based peer assessment system. *Canadian Journal of Learning and Technology*, 34(2). Retrieved from <http://www.cjlt.ca/index.php/cjlt/article/viewArticle/497/228>
- Ozden, Y., Erturk, I., & Sanli, R. (2004). Student perceptions of online assessment: A case study. *Journal of Distance Education*, 19(2), 66-92.
- Walker, D., Topping, K., & Rodrigues, S. (2008). Student reflections on formative e-assessment: expectations and perceptions. *Learning, Media and Technology*, 33(3), 221-234. doi:10.1080/17439880802324178
- Yates, R., & Beaudrie, B. (2009). The impact of online assessment on grades in community college distance education courses. *The American Journal of Distance Education*, 23(2), 62-70. doi: 10.1080/08923640902850601
- Yilmaz, H. (2010). The implementation and evaluation of a peer assessment based online course. *Journal of Instructional Technology and Distance Learning*, 4(5). Retrieved from [http://itdl.org/Journal/Apr\\_10/article05.htm](http://itdl.org/Journal/Apr_10/article05.htm)
- Zimmer, L. (2004). Qualitative meta-synthesis: a question of dialoguing with texts. *Journal of Advanced Learning*, 53(3), 311-318. Retrieved from [http://www.hkr.se/upload/HS/doc/Forskningsplattform\\_24april\\_2.pdf](http://www.hkr.se/upload/HS/doc/Forskningsplattform_24april_2.pdf)