



Hacettepe University Journal of Education  
Hacettepe Üniversitesi Eğitim Fakültesi Dergisi  
e-ISSN: 2536-4758



E-Sınavlardan Etkin Geri Bildirim Üretmek için Öneriler\*

Okan BULUT\*\*, Maria CUTUMISU\*\*\*, Deepak SINGH\*\*\*\*, Alexandra M. AQUILINA\*\*\*\*\*

Makale Bilgisi	ÖZET
<b>Geliş Tarihi:</b> 10.05.2020	<p>Günümüzün eğitim sistemleri öğrencilerin öğrenme sürecini desteklemek amacıyla birçok yeni teknolojiyi uyarlamaya ve kullanmaya devam etmektedir. Bu teknolojilerden birisi de e-sınavlardır. Bu tarz sınavlar öğrencilerin sınav sorularını bilgisayar ya da tablet gibi dijital araçlar ile cevaplamasına imkân vermektedir. E-sınavların önemli faydalarından biri de etkili, hızlı ve bireye özgü geri bildirim üretmeye imkân sağlamasıdır. Bununla beraber e-sınavlardan etkin olarak geri bildirim üretilmesine dair önerilerin yer aldığı çalışmalara literatürde rastlanılmamıştır. Literatürdeki bu önemli boşluğu doldurmak adına e-sınavlarda geri bildirim üretimine dair sistematik bir derleme çalışması yapılmıştır. Bu çalışma üç aşamada gerçekleştirilmiştir. İlk aşamada geri bildirim üretimine dair uygulamaları içeren çalışmalar derlenmiş ve önemli noktalar özetlenmiştir. İkinci aşamada ise e-sınavlardan etkin geribildirim üretmek için araştırmacı ve uygulayıcılara yönelik öneriler sunulmuştur. Son aşamada ise e-sınavlardan hızlı ve bireye özgü geri bildirim üretiminde kullanılabilir altı adımdan oluşan bir yönerge ortaya konmuştur. Bu yönergede görsel ve yazılı öğelerden oluşan geri bildirimler aracılığıyla öğrencilerin sınav sonuçlarını nasıl daha etkin bir şekilde yorumlayabileceği gösterilmiştir.</p> <p><b>Keywords:</b> E-sınav, geri bildirim, sınav raporu, sınav geliştirme, bilgisayar destekli sınav</p>
<b>Kabul Tarihi:</b> 29.06.2020	
<b>Erken Görünüm Tarihi:</b> 29.09.2020	
<b>Basım Tarihi:</b> 30.09.2020	

Guidelines for Generating Effective Feedback from E-Assessments

Article Information	ABSTRACT
<b>Received:</b> 10.05.2020	<p>Today's education systems continue to adopt new technologies to support student learning. One of these technologies is e-assessment, a form of assessment that enables students to answer items using digital devices, such as computers and tablets. One of the benefits of e-assessments is the ability to generate interactive, timely, and customized feedback for students. Yet, despite vast literature on the generation and delivery of feedback, there is no systematic review of the guidelines on how e-assessments can be used for generating effective feedback. The objectives of this study are threefold. First, we synthesize the literature on the current practices in feedback generation. Second, we provide researchers and practitioners with a synthesis of guidelines for best practices in generating effective feedback with e-assessments. Third, we introduce a new framework in which we demonstrate the six steps of creating an e-assessment that can help produce immediate, customized, and specific feedback for students. This framework combines multiple forms of feedback (e.g., graphs, tables, and text) to improve the understanding of feedback and engage students in the interpretation of their score reports. Implications for practice and future research are discussed.</p> <p><b>Keywords:</b> E-assessment, feedback, score reporting, test development, computerized assessment</p>
<b>Accepted:</b> 29.06.2020	
<b>Online First:</b> 29.09.2020	
<b>Published:</b> 30.09.2020	

doi: 10.16986/HUJE.2020063705 Makale Türü (Article Type): Research Article

**Kaynakça Gösterimi:** Bulut, O., Cutumisu, M., Singh, D., & Aquilina, A. M. (2020). E-sınavlardan etkin geri bildirim üretmek için öneriler. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 35(Özel Sayı), 60-72. doi: 10.16986/HUJE.2020063705

**Citation Information:** Bulut, O., Cutumisu, M., Singh, D., & Aquilina, A. M. (2020). Guidelines for generating effective feedback from e-assessments. *Hacettepe University Journal of Education*, 35(Special Issue), 60-72. doi: 10.16986/HUJE.2020063705

\* This work was supported by a Social Sciences and Humanities Research Council of Canada (SSHRC) Insight Development Grant (430-2016-00039) awarded to Dr. Okan Bulut.

\*\* Associate Professor, University of Alberta, Faculty of Education, Centre for Research in Applied Measurement and Evaluation, Edmonton-CANADA. e-mail: [bulut@ualberta.ca](mailto:bulut@ualberta.ca) (ORCID: 0000-0001-5853-1267)

\*\*\* Associate Professor, University of Alberta, Faculty of Education, Centre for Research in Applied Measurement and Evaluation, Edmonton-CANADA. e-mail: [cutumisu@ualberta.ca](mailto:cutumisu@ualberta.ca) (ORCID: 0000-0003-2475-9647)

\*\*\*\* Graduate Research Assistant, University of Alberta, Faculty of Education, Department of Educational Psychology, Edmonton-CANADA. e-mail: [dsingh1@ualberta.ca](mailto:dsingh1@ualberta.ca) (ORCID: 0000-0003-3500-5314)

\*\*\*\*\* Graduate Research Assistant, University of Alberta, Faculty of Education, Department of Educational Psychology, Edmonton-CANADA. e-mail: [aquilina@ualberta.ca](mailto:aquilina@ualberta.ca) (ORCID: 0000-0003-1876-6015)

## 1. INTRODUCTION

Many researchers have portrayed feedback as an important catalyst for improving student learning and academic performance (Bailey & Garner, 2010; Evans, 2013; Hattie & Gan, 2011). As an essential component of both summative and formative assessments, feedback serves several functions including providing students with information about the accuracy of their responses, the knowledge and understanding needed for correct responses, and helping students acquire essential knowledge (Narciss, 2008; Shute, 2008). Formative feedback generated by teachers enables students to capitalize on their strengths and identify their weaknesses, as well as guide them towards the necessary steps required to achieve the learning outcomes (Hatziapostolou & Paraskakis, 2010). Furthermore, acknowledging students' success and academic progress through feedback helps students develop a positive attitude toward learning and motivates them to learn more (Daniels & Bulut, 2019; Yuan & Kim, 2015).

In today's ever-changing technology landscape, education systems continue to evolve and adopt new technologies to support student learning. For example, e-assessment tools such as computer-based and computerized adaptive tests can be used to provide students with immediate feedback enriched with visual and interactive elements (e.g., Bulut, Cutumisu, Aquilina, & Singh, 2019). Furthermore, intelligent tutoring systems enable automated and adaptive feedback generation for learners within a personalized learning environment (e.g., Gutierrez & Atkinson, 2011). Similarly, massive open online courses (MOOCs) create additional learning opportunities, such as receiving feedback through online peer assessment (Suen, 2014). Despite these innovative approaches, generating effective feedback from e-assessments is still considered challenging because students are more likely to ignore feedback messages in e-assessment settings due to the lack of face-to-face interaction with their teacher (Wuensch, Aziz, Ozan, Kishore & Tabrizi, 2008; Timmers & Veldkamp, 2011; Van der Kleij, Eggen, Timmers & Veldkamp, 2012). Furthermore, students do not necessarily review their feedback carefully and even if they do, they may not understand or use it to enhance their learning (Prince, Handley, Millar, O'Donovan, 2010).

As emerging technologies continue to enable instructors to create high-quality assessments, incorporating effective feedback practices into e-assessments becomes highly essential. The goals of the current study are threefold. First, we will synthesize the literature on the current practices in feedback generation. Second, we will provide researchers and practitioners with a set of guidelines regarding the generation and delivery of effective feedback using e-assessments. Third, we will present a new framework in which we demonstrate the steps of creating an e-assessment that can help generate immediate, customized, and specific feedback for students. With this framework, instructors can design a fine-grained e-assessment that not only improves the content validity but also facilitates the generation and delivery of personalized feedback.

## 2. THE LITERATURE ON FEEDBACK GENERATION

### 2.1. Qualities of Effective Feedback

Previous research suggests that effective feedback must have several important qualities to promote learning and improve achievement. To be effective, feedback must be specific, timely, understandable, non-threatening, and revisable (STUNR; Schwartz, Tsang, & Blair, 2016). The following section will summarize each of the STUNR principles as well as other important characteristics of feedback.

#### 2.1.1. Specific

For feedback to be effective, it must be content-specific, either goal or task-oriented (Hattie & Timperley, 2007), and neutral (Thurlings, Vermeulen, Bastiaens, & Stijnen, 2013). Feedback should help students understand their learning goals, self-monitor their progress, and identify ways to improve their performance (Yuan & Kim, 2015). Elaborated feedback (e.g., providing a directive explanation) is significantly more effective than feedback indicating whether or not the answer was correct or simply providing the correct answer (Schartel, 2012). This is particularly the case for higher-order learning outcomes (Van der Kleij, Feskens, & Eggen, 2015). Generally, elaborated and specific feedback is perceived as more useful than feedback that is general and brief (Harks, Rakoczy, Hattie, Besser, & Klieme, 2014). Moreover, feedback that simply praises the student fails to improve learning outcomes (Hattie & Gan, 2011; Hattie & Timperley, 2007), although this type of feedback may have a positive impact on students' motivation and perseverance (Van der Kleij et al., 2015). However, it is important to ensure that feedback is manageable, since overly detailed feedback with countless comments may become confusing for some students who may struggle with distinguishing important comments from the less important ones (Race, 2006).

#### 2.1.2. Timely

The literature on the timing of feedback generally displays conflicting results, as many factors may influence this aspect of feedback delivery. However, many researchers suggest that students generally prefer immediate feedback to delayed feedback (Epstein et al., 2002; Daniels & Bulut, 2019; Yuan & Kim, 2015). Feedback seems to be more effective if it is provided quickly, while students can still remember how they addressed each assessed task (Race, 2006). However, the timing of feedback is dependent on the type of skill as well as on the level of task difficulty (Fluckiger, Vigil, Pasco, & Danielson, 2010). For example, during fluency tasks, immediate corrective feedback may have a negative impact on students (Hattie & Timperley, 2007). In

comparison, during difficult tasks delayed feedback may be useful, as it provides students with an opportunity to process the information related to the task (Clariana, Wagner, & Murphy, 2000). Shute (2008) suggested that for lower-order learning outcomes, immediate feedback works best, but for higher-order learning outcomes, delayed feedback works more effectively. Other research has also demonstrated that delayed feedback could lead to superior final test performance in comparison to immediate feedback, due to the spaced presentation of information (Butler, Karpicke, & Roediger, 2007). Also, the timing of feedback could depend on the source of feedback (i.e., teacher or peer). For example, Kulkarni, Bernstein, and Klemmer (2015) found that peer feedback becomes more effective when it is delivered within 24 hours and especially within an hour.

### ***2.1.3. Individualized, non-threatening, and supportive of individual growth***

In addition to being timely, effective feedback should also support individual development and inform students about both their strengths and their weaknesses (Lilley & Barker, 2007). Feedback must be individualized, non-evaluative, and supportive. It should also address what the students did well, what areas they need to improve, and how they can improve their performance (Desrochers & Zell, 2012; Jonsson, 2012; Schartel, 2012; Shute, 2008). Feedback can promote learning when it is personal, manageable, motivational, and directly associated with the assessment guidelines and learning outcomes (Hatzia Apostolou & Paraskakis, 2010). To increase student motivation and encourage students to perform better, feedback must be constructive and empowering. Generally, students prefer feedback that is provided to them in person as it allows for an opportunity for discussion with the teacher and identifies areas in need of further improvement (Beaumont, O'Doherty, & Shannon, 2011; Blair, Curtis, Goodwin, & Shields, 2013).

### ***2.1.4. Comprehensible***

For feedback to be effective, students must be able to comprehend the content of feedback. Thus, feedback must be clear and congruent with the knowledge level of students. Furthermore, a dialogue about the comments with the teacher allows students to reflect on their feedback and check their understanding of feedback (Yuan & Kim, 2015). Some students have also shown a preference for written feedback, as it provides them with an opportunity to reflect on the assignment at a later date (Blair et al., 2013).

### ***2.1.5. Actionable***

Effective feedback incorporates strategies to ensure that students read and use feedback to improve their future performance. These strategies include revising and resubmitting work as a result of feedback, as well as writing a summary of the students' changes to demonstrate how they used feedback (Yuan & Kim, 2015). Furthermore, feedback should lead to a collaboration between the teacher and the students, be based on first-hand data, be restricted to behaviors that can be changed, and deal with decisions and actions rather than presumed intentions or interpretations (Schartel, 2012). Daniels and Bulut (2019) suggested that providing students with effective feedback may not be enough to influence students' subsequent performance and thus instructors should also help students build a plan of action based on the received feedback.

## **2.2. Lessons Learned from Feedback Research**

Previous studies on feedback have highlighted many challenges in the generation, delivery, and use of feedback in practice. Studies show that there is often a big gap between teachers' expectations of the impact of feedback on student performance and the perceived usefulness of feedback among students. Some of the potential causes of this gap include failing to understand feedback (Lea & Street, 1998), prioritizing grades over feedback (Black & Wiliam, 1998), and lacking an action plan based on feedback (Daniels & Bulut, 2019). Also, students are usually not trained or supported in their use of feedback (Carless, Salter, Yang, & Lam, 2011; Quinton & Smallbone, 2010). Therefore, providing students with guidance on the content and interpretation of feedback could be a key factor influencing the effectiveness of feedback for students.

Research indicates that feedback is not a one-way interaction between the teacher and students. Instead, it is a part of a larger learning ecosystem that includes two-way interactions between many elements, such as the teacher, students, peers, the learning environment, and additional resources. When attempting to generate effective feedback for students, it is important to take into account the interaction between student and teacher mediators within the learning ecosystem. Some of these mediators include intelligence, personality, motivation, gender, culture, cognitive styles, educational experiences, and beliefs about learning (Evans, 2013). Evans (2013) also argued that the role, interrelationships, and importance of these mediators are likely to change over time.

Another important finding from previous research is that the feedback that students prefer to receive and the feedback that would benefit them are not necessarily the same (Jonsson, 2012). Race (2006) argued that, contrary to popular belief among students, the optimal type of feedback may not be specific, detailed, positive, and individualized. The feedback that is less specific and individualized could encourage students to seek further clarification on the provided information and lead to a productive learning experience (Jonsson, 2012). In line with these findings, Bulut et al. (2019) found that a majority of undergraduate students who reported specific and detailed feedback as their preference were not willing to review their exam reports that

included specific and detailed feedback on their performance. Instead, most students preferred to review a brief report with general feedback that was provided upon completion of the exam.

### **3. FEEDBACK IN E-ASSESSMENTS**

Computerized feedback (i.e., computer-provided feedback) is a recently developed method of providing students with feedback, particularly in response to e-assessments. To date, researchers have reported both positive and negative findings regarding the effects and utilization of computerized feedback in education. Therefore, before deciding whether or not to employ computerized feedback, it is important to consider its advantages and disadvantages in practice. The following sections provide a summary of the advantages and disadvantages of computerized feedback, followed by a set of guidelines for generating effective feedback from e-assessments. Finally, a new framework for generating effective feedback from e-assessments is introduced.

#### **3.1. Advantages of Computerized Feedback**

##### **3.1.1. Timing**

Students participating in e-assessments can receive feedback more rapidly because computerized feedback is often automatically generated based on students' responses to the items (Epstein et al., 2002; Yuan & Kim, 2015). For example, the e-assessment systems are capable of providing immediate feedback through an answer-until-correct option in which students have the opportunity to make multiple attempts until they find the correct answer. This approach promotes the acquisition and retention of knowledge (Epstein et al., 2002). Also, immediate feedback in an e-assessment setting leads to a positive impact on learning, as students are still able to recall how they addressed each task, which would not be possible in a paper-pencil testing setting (Race, 2006; Van der Kleij et al., 2015). In K-12 large-scale assessments, digital score reports including feedback can be shared rapidly with students, parents, schools, and other stakeholders (Desrochers & Zell, 2012; Kyllonen, 2009). Furthermore, the availability of immediate computerized feedback reduces the workload for teachers, especially when large numbers of students or long tests are involved (Timmers, 2013).

##### **3.1.2. Diversity**

The dynamic nature of e-assessments enables greater diversity in computerized feedback with regard to the content, type, and amount of information being presented to students. Research shows that, in addition to conventional forms of feedback, computerized feedback can be provided in many other forms, such as affective and emotional feedback (Moridis & Economides, 2012), positive and negative feedback (Weedon, 2000), individualized feedback (Wu, Kuo, & Wang, 2017), and peer-to-peer feedback (McCarthy, 2017). Also, students can be given the option to choose what type of feedback (e.g., corrective feedback or diagnostic feedback) they might prefer before computerized feedback is generated from their responses to an e-assessment.

##### **3.1.3. Diagnostic information**

Computerized feedback is helpful for students to understand how successful they were in their learning (Lilley & Barker, 2007). In addition to indicating students' overall performance, computerized feedback can also be used for generating more specific and diagnostic information from e-assessments. When e-assessments are used as a formative assessment tool, computerized feedback generated from these assessments can help diagnose students' conceptual understanding of declarative knowledge, such as science education (Maier, Wolf, & Randler, 2016). Furthermore, the use of computerized feedback positively impacts the learning experience of students by guiding their efforts in learning and diagnosing misconceptions as well as areas of difficulty in their learning (Fui & Lian, 2018; Lowry, 2005).

#### **3.2. Disadvantages of Computerized Feedback**

##### **3.2.1. Perceived usefulness**

Generally, teachers anticipate that all students will perceive feedback in the way they intended it to be perceived (Harks et al., 2014). However, researchers argued that computerized feedback might not be perceived as useful as feedback given by teachers because students would be less likely to accept comments, praise, or criticism delivered via computers (e.g., Budge, 2011; Lepper, Woolverton, Mumme, & Gurtner, 1993). These studies indicate that students tend to use teacher-delivered feedback as a baseline when evaluating the effectiveness of computerized feedback. In an empirical study with college students, Lipnevich and Smith (2008) found that students rated the instructor's feedback as being more accurate and helpful than computerized feedback, despite the common belief that computerized feedback would be more trusted due to computers' capability to generate more neutral and unbiased information.

### 3.2.2. Disregard

Previous research indicated that it can be more challenging to provide students with computerized feedback in an e-assessment setting because it is often easier for students to ignore computer-delivered feedback messages (Wuensch et al., 2008; Timmers & Veldkamp, 2011; Van der Kleij et al., 2012). Due to the lack of face-to-face interaction with the instructors, students do not necessarily read their computerized feedback and, if they do, they may not understand or use it (Prince et al., 2010).

### 3.2.3. No follow-up

In an e-assessment setting, there is limited ability for students to engage in additional communication or follow-up with teachers regarding their feedback (Yuan & Kim, 2015). Race (2001) argued that students who fail to understand computerized feedback may not be able to follow up with their teacher to ask for further clarification, as they would normally do after receiving instructor-delivered feedback. Therefore, the impact of computerized feedback on student learning could be less positive than anticipated.

## 3.3. Guidelines for e-assessments

During the last two decades, there have been many publications, presentations, and technical reports including guidelines for the generation and delivery of feedback (e.g., Bulut, Cutumisu, Singh, & Aquilina, 2018; Jug, Jiang, & Bean, 2019; Nicol & Macfarlane-Dick, 2006; Pendleton, Schofield, Tate, & Havelock, 2003; Slater, Livingston, & Silver, 2019; Zapata-Rivera & Katz, 2014; Zenisky & Hambleton, 2012, 2016). However, these guidelines do not specifically focus on e-assessments. Therefore, a review of the existing resources that contain the current guidelines for generating effective feedback from e-assessments is needed. This section aims to provide a brief synthesis of guidelines for best practices in generating effective feedback with e-assessments. The guidelines were grouped into three categories: test development, the content of feedback, and the delivery of feedback. Using these categories, Table 1 presents a list of guidelines found in the feedback and score reporting literature.

Table 1.

#### *The Guidelines for Generating Effective Feedback with e-assessments*

Category	Guidelines
Test development	<p>The assessment should be deliberately designed to improve student performance (Wiggins, 1998).</p> <p>The assessment should contain higher-order questions that are clearly worded and transparent in their marking (Walker, Topping, &amp; Rodrigues, 2008).</p> <p>The assessment should include clear instructions on how to answer each item (Walker et al., 2008).</p> <p>The assessment should have a formative function providing feedforward for future learning (Bloxham, 2015, p. 109).</p> <p>If possible, the instructor should allow students to engage in test construction (Nicol, 2007).</p> <p>The teacher (or test developer) should create a test blueprint linking the items to the feedback generation process (Bulut et al., 2018).</p> <p>If subscores by content categories are to be provided, each content category should have enough items with distinct information to produce reliable subscores (Bulut et al., 2018; Bulut, Davison, &amp; Rodriguez, 2017; Sinharay, 2010).</p>
Content of feedback	<p>The score report should be tailored to meet the needs and characteristics of the target audience, such as students, parents, and teachers (Hambleton &amp; Zenisky, 2013; Zapata-Rivera, &amp; Katz, 2014).</p> <p>The score report should have an aesthetically pleasing design without information overload (Bulut et al., 2018; Slater et al., 2019).</p> <p>The score report should present the feedback in different forms, including narrative text, tables, and figures (Bulut et al., 2018).</p> <p>The layout of the score report should be simple, with key results highlighted (Goodman &amp; Hambleton, 2004; Slater et al., 2019).</p> <p>Feedback presented in the score report should include a set of actions that students can take to improve their future performance (Daniels &amp; Bulut, 2019; Hattie, 2009; Jonsson, 2012).</p> <p>If interactive elements (e.g., visuals and tables) are to be used, how students will interact with these elements should be considered in the design process (Bulut et al., 2019; Slater et al., 2019).</p> <p>Usability studies with students should be carried out to test whether the content of feedback is easy to follow (Slater et al., 2019; Zenisky &amp; Hambleton, 2012).</p>
Delivery of feedback	<p>If there is an online location (e.g., a website) to view score reports, the homepage design should consider student interest and needs (Zenisky &amp; Hambleton, 2012).</p> <p>If online score reporting is used, downloadable PDFs that print out the same information in a clear way should also be included (Zenisky &amp; Hambleton, 2012).</p> <p>Students should be given the option to choose the timing (i.e., immediate or delayed) of feedback (Bulut et al., 2018).</p> <p>Frequency of feedback (e.g., after each item or at the end) should be determined based on the type of e-assessment (i.e., summative or formative) and the number of items on the e-assessment (Bulut et al., 2018).</p>

### 3.4. A New Framework for Generating Feedback through e-Assessments

In this study, we also introduce a new framework that can be used for generating effective feedback from e-assessments. The main objective of the proposed framework is to encourage instructors and test developers to design a balanced assessment in terms of content domains, concepts targeted within each domain, and the number of items. The framework involves six steps. Figure 1 depicts each of these steps needed for generating effective feedback.

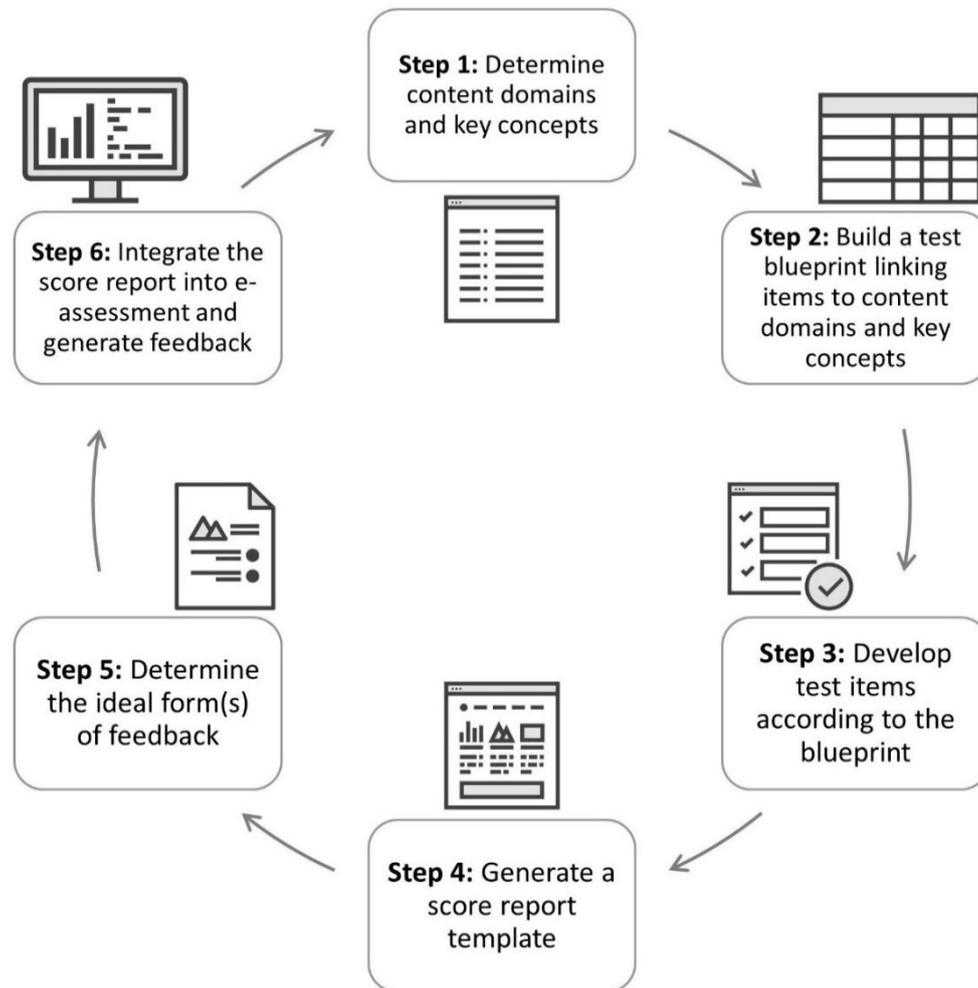


Figure 1. The workflow of the feedback generation process for e-assessments

In the first step, the instructor should determine the target content domains that need to be covered in the assessment. Also, the instructor should identify the key concepts within each content domain. The key concept refers to specific content knowledge or skills that students are expected to gain as they learn about each content domain. This step is highly crucial for the content validity of the assessment because drawing meaningful inferences from the assessment results depends on the degree to which the assessment measures the target construct adequately (Messick, 1989; Moss, 1995).

The second step focuses on building a test blueprint for the assessment. A test blueprint, also known as the table of test specifications, is essentially a two-way table of content domains (rows) and cognitive complexity levels (columns) intended to be included in the assessment. Haladyna and Rodriguez (2013) defined cognitive complexity as “the expected mental complexity involved when a test item is administered to a typical test taker” (p. 28). The test blueprint delineates the number (or proportion) of items for each combination of the content domains and cognitive complexity levels. The number of items reflects the relative weight of each content domain in the assessment. The total number of items for each row indicates the relative importance of each content domain, while the total number of items specified for each column represents the relative distribution for each level of cognitive complexity, such as the levels of remembering, understanding, and applying in Bloom’s Revised Taxonomy (Anderson et al., 2001).

According to Perie and Huff (2015), what students should know and be able to do is one of the major questions that should drive the development of test blueprints for assessments. The test blueprint can guide the instructor significantly during the test development process, especially for item writing and test assembly. The instructor can develop a balanced assessment according to the test blueprint to ensure that the assessment appropriately reflects the content domains and the intended levels of cognitive complexity. However, using a traditional test blueprint, the instructor may still struggle with identifying what specific knowledge or skills need to be measured within each content domain. Therefore, the test blueprint could be expanded

with an additional column that provides a listing of the key concepts for each content domain. During the test development process, the instructor must ensure that each item on the assessment is linked to at least one of the key concepts listed in the test blueprint. Important key concepts could be targeted by multiple items in the assessment. Figure 2 shows an example of a test blueprint for a hypothetical assessment that is planned to include 40 items from four content domains. Each content domain has several key concepts that are expected to guide the instructor when creating and assembling items for the assessment (see Bulut et al., 2019 for the examples of key concepts created for an undergraduate course).

Content	Key Concepts	Remembering	Understanding	Applying	Total	%
A	Key concept 1, Key concept 2, Key concept 3, ...	8	4	4	16	40%
B	Key concept 1, Key concept 2, Key concept 3, ...	2	6	-	8	20%
C	Key concept 1, Key concept 2, Key concept 3, ...	4	4	4	12	30%
D	Key concept 1, Key concept 2, Key concept 3, ...	-	-	4	4	10%
<b>Total</b>		14	14	12	40	
<b>%</b>		35%	35%	30%		

Figure 2. A test blueprint for a hypothetical assessment with 40 items

The third step focuses on the development of test items according to the test blueprint. Unless it is a new assessment, the instructor can utilize previously developed items to build the assessment according to the test blueprint. If, however, the instructor needs to create new items, then the key concepts specified in the test blueprint could guide the item writing process. The instructor also needs to determine item format (e.g., selected-response and constructed-response), depending on the content domains, target levels of cognitive complexity, and the student population (see Haladyna and Rodriguez, 2013 for a comprehensive review of the item and test development processes). The final product of the item writing process should be an expanded, item-level form of the test blueprint in which each item is categorized based on their content domain, key concepts, and cognitive complexity. Figure 3 illustrates an expanded test blueprint using the same example of a hypothetical assessment with 40 items.

In the fourth step, a score report template needs to be created. The template can be created for either dynamic reporting (e.g., a web-based report with interactive features) or static reporting (e.g., a static document to download or print). The template should include the total score (e.g., total raw score, percent-correct score, or scaled score) as well as scores by content categories (i.e., subscores). Subscores can provide students with more fine-grained information from the assessment beyond a total score and help them identify their strengths and weaknesses more clearly (Bulut et al., 2017). In addition to the scores, the score report can also include other elements, such as narrative text, graphics, and tables. To create an aesthetically pleasing report for students, the principles of visual design (e.g., scale, contrast, balance, and visual hierarchy) and cognitive psychology should be considered during the design process. These design principles not only facilitate readability and interpretation of the feedback provided in the score reports, but also can draw the student's attention to the important information presented in the report (Gotch & Roduta Roberts, 2018).

Item	Item Description	Content Domain	Key Concept	Cognitive Complexity
1	...	A	Key concept 1	Remembering
2	...	B	Key concept 3	Applying
3	...	C	Key concept 1	Understanding
.		.	.	.
.		.	.	.
.		.	.	.
40	...	B	Key concept 2	Remembering

Figure 3. An expanded form of the test blueprint after the test development process

In the fifth step, the instructor needs to determine the optimal ways to present feedback to the students. A score report typically includes a combination of interactive or static graphics, tables, numbers, and narrative text. The score report can present the feedback in a simple format (e.g., mostly narrative text and numbers) or in a relatively more advanced format (e.g., interactive graphics and tables). However, a balanced design can be developed by including multiple ways to present feedback and considering the target audience. For example, Bulut et al. (2019) found that undergraduate students did not prefer score reports that present feedback in a single format (e.g., only graphics or tables). Instead, most students wanted to receive their feedback in multiple formats, such as narrative text, graphics, and tables.

In the final step, the score report template is integrated into the e-assessment and used for generating feedback as students take the assessment. Unless the assessment contains constructed-response items that require human scoring, the score reports can be automatically generated immediately after the students complete the assessment. However, previous research shows that providing immediate feedback in the e-assessment settings could lead to negative consequences, such as avoiding reviewing the score reports after the assessment period is over (Bulut et al., 2019; Van der Kleij, 2012) and failing to understand and process the feedback (Prince et al., 2010). Therefore, the instructor may consider making the score reports available to their students once the students recover from mental fatigue that occurs due to prolonged cognitive activity when completing the assessment.

The feedback generation framework introduced in this study has several implications for researchers and practitioners. First, the framework allows researchers to investigate the effects of many elements and conditions in feedback generation, such as the timing of feedback (e.g., immediate or delayed), type of feedback (e.g., graphics, tables, or text), and density of feedback (e.g., concise or detailed). Second, this framework could be possibly extended to a small classroom assessment where the instructor uses both multiple-choice and constructed-response items together and scores the items manually. Following the steps of feedback generation (see Figure 6), the instructor can build a template using a particular software program (e.g., Microsoft Excel) and create individual score reports with different types of feedback. Third, the feedback generation framework can be used with other innovative applications (e.g., automatic item generation, automated test assembly, and automated score reporting) to build a more efficient and effective assessment platform for large-scale testing.

#### 4. DISCUSSION

Today's education systems continue to evolve and adopt new technologies to support and enhance student learning. One of these technologies is e-assessment that refers to the use of information technology in conducting assessments (Singh & Villiers, 2017). Using e-assessment tools (e.g., computers, tablets, and smartphones), students can answer items in a digital environment. Previous research suggests that e-assessments facilitate the generation and delivery of effective feedback while students are solving the items. The instructor can use e-assessments to provide students with feedback that is not only timely but also tailored to students' needs (van der Kleij et al., 2012; Lopez, 2009). However, the existing literature does not present consistent evidence regarding the conditions in which timely and customized feedback generated from e-assessments can contribute to student learning. Therefore, this study aimed to present a review of the feedback literature over the past two decades, with a particular emphasis on the generation and delivery of feedback with e-assessments. The lessons learned from the feedback literature were discussed. Then, a review of the published feedback guidelines related to e-assessments was presented. Finally, a new framework for creating an e-assessment that can help generate immediate, customized, and specific feedback for students was introduced.

As Hepplestone, Holden, Irwin, Parkin, and Thorpe (2011) also indicated, there is a vast literature on effective feedback practices in education, whereas the literature on the use of technology to enhance the production, delivery, and utilization of feedback is still limited. The literature on the characteristics of effective feedback suggests that feedback must be specific, timely, easy to understand, non-threatening, revisable, and actionable (Jonsson, 2012; Schwartz et al., 2016; Shute, 2008). However, researchers also pointed out that there is a complex relationship among the effectiveness of feedback, student characteristics (e.g., personality, intelligence, and motivation), and other potential mediators (e.g., instructor, peers, and learning environment) in practice (Evans, 2013; Shute, 2008). In addition, research suggests that the perceived usefulness of feedback among students

depends on many factors, such as the content of feedback, student motivation, and students' familiarity with feedback (Bulut et al., 2019; Carless et al., 2011; Daniels & Bulut, 2019; Evans, 2013).

Recent studies appear to support the hypothesis that the quality, quantity, and use of feedback can be enhanced with the use of technology in education. Students participating in e-assessments can receive immediate feedback upon completion of the assessment, which would not be possible traditional paper-and-pencil testing (Van der Kleij et al., 2015; Yuan & Kim, 2015). The automatic generation of feedback from e-assessments could reduce the workload for teachers who are expected to evaluate a large group of students (Timmers, 2013). Feedback provided through online score reporting also offers a certain level of flexibility in accessing and reviewing the comments at a convenient time for each student (Hepplestone et al., 2011). In addition to the students, online score reports including feedback can also be shared with parents, schools, and other stakeholders (Desrochers & Zell, 2012; Kyllonen, 2009).

Despite the benefits of technology in supporting the generation and delivery of feedback, researchers also highlighted potential disadvantages of computerized feedback. For example, Budge (2011) found that students are less likely to accept comments, praise, or criticism provided within an e-assessment setting. Similarly, Lipnevich and Smith (2008) argued that teacher-delivered feedback is often perceived as more useful by students. Another potential challenge is that students are less likely to follow up with their instructors to seek further clarification on their feedback. Due to the lack of interaction with their teachers, students tend to disregard computer-delivered messages (Wuensch et al., 2008; Timmers & Veldkamp, 2011).

The existing guidelines found in the feedback literature focus on three aspects of the feedback generation process: test development, feedback content creation, and feedback delivery. For a balanced assessment that can produce effective feedback, the use of a test blueprint in the test development process appears to be essential (Bulut et al., 2018). In addition, instructors are recommended to consider a formative function for the e-assessment, so it can be used for improving student performance in future assessments. Regarding the content of feedback, researchers highlighted the necessity of utilizing the principles of visual design in creating a score report that is simple, easy to navigate, and aesthetically pleasing (Bulut et al., 2018; Slater et al., 2019; Zenisky & Hambleton, 2012). Researchers also made several suggestions related to the delivery of feedback. When determining the timing and frequency of feedback, the type of assessment, the length of assessment, and the characteristics of the target student group should be considered (Bulut et al., 2018).

Although the published guidelines and recommendations on the generation and delivery of feedback aim to enhance the effectiveness of feedback practices in the classroom, it should be acknowledged that some of these guidelines are not based on empirical evidence. With this limitation in mind, this study offers some insights into the conditions in which e-assessments can produce effective feedback that can support student learning. The feedback generation framework introduced in this study follows a teacher-centered approach in the design of both the assessment and feedback to be generated from the assessment. As instructors attempt to build a reliable assessment that can yield valid conclusions about the students, feedback generation should also be considered a key component, not an afterthought. Furthermore, in order to provide effective feedback to the students, instructors must employ the principles of both visual design and communication when designing score reports. Future research should focus on the complex interaction between student characteristics (e.g., age, grade level, and motivation level) and the feedback generated with e-assessments to better guide the instructors and test developers.

### **Research and Publication Ethics Statement**

This study was carried out in accordance with the recommendations of the Research Ethics Office and the protocol (Pro00072612) was approved by the Research Ethics Board 2 at the University of Alberta.

### **Contribution Rates of Authors to the Article**

Research idea: OB (60%), MC (40%); Literature review: OB (25%), MC (25%), DS (25%), AA (25%); Manuscript preparation: OB (60%), MC (20%), DS (10%), AA (10%).

### **Support Statement**

This work was supported by a Social Sciences and Humanities Research Council of Canada (SSHRC) Insight Development Grant (430-2016-00039) awarded to Dr. Okan Bulut.

### **Statement of Interest**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## 5. REFERENCES

- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Pearson: Allyn & Bacon.
- Bailey, R., & Garner, M. (2010). Is the feedback in higher education assessment worth the paper it is written on? Teachers' reflections on their practices. *Teaching in Higher Education*, 15(2), 187–198.
- Beaumont, C., O'Doherty, M., & Shannon, L. (2011). Reconceptualising assessment feedback: a key to improving student learning? *Studies in Higher Education*, 36(6), 671–687.
- Black, P. & Wiliam, D. (1998) Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Blair, A., Curtis, S., Goodwin, M., & Shields, S. (2013). What feedback do students want? *Politics*, 33(1), 66–79.
- Bloxham, S. (2015). Assessing assessment: New developments in assessment design, feedback practices and marking in higher education. In H. Fry, S. Ketteridge, & S. Marshall (Eds.), *A handbook for teaching and learning in higher education* (4<sup>th</sup> ed., pp. 107–122). Abingdon: Routledge.
- Budge, K. (2011). A desire for the personal: student perceptions of electronic feedback. *International Journal of Teaching and Learning in Higher Education*, 23(3), 342–349.
- Bulut, O., Cutumisu, M., Aquilina, A. M., & Singh, D. (2019). Effects of digital score reporting and feedback on students' learning in higher education. *Frontiers in Education*, 4(65), 1–16. doi: 10.3389/educ.2019.00065
- Bulut, O., Cutumisu, M., Singh, D., & Aquilina, A. (2018, July). *Guidelines for generating effective feedback from computer-based assessments*. Paper presented at the International Testing Commission Conference, Montreal, QC, Canada.
- Bulut, O., Davison, M. L., & Rodriguez, M. C. (2017). Estimating between-person and within-person subscore reliability with profile analysis. *Multivariate Behavioral Research*, 52(1), 86–104. doi: 10.1080/00273171.2016.1253452
- Butler, A. C., Karpicke, J. D., & Roediger III, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, 13(4), 273–281.
- Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, 36(4), 395–407.
- Clariana, R. B., Wagner, D., & Murphy, L. C. R. (2000). Applying a connectionist description of feedback timing. *Educational Technology Research and Development*, 48, 5–22.
- Daniels, L. M., & Bulut, O. (2019). Students' perceived usefulness of computerized percentage-only vs. descriptive score reports: Associations with motivation and grades. *Journal of Computer Assisted Learning*, 36(2), 199–208. doi: 10.1111/jcal.12398
- Desrochers, C. G., & Zell, D. (2012). *Provided timely and frequent feedback on tests, reports, projects, etc., to help students improve*. IDEA Paper No. 17. Manhattan, KS: The IDEA Center.
- Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, 83(1), 70–120.
- Epstein, M. L., Lazarus A. D., Calvano, T. B., Matthews, K. A., Hendel, R. A., Epstein, B. B., & Brosvic, G. M. (2002). Immediate feedback assessment technique promotes learning and corrects inaccurate first responses. *The Psychological Record*, 52, 187–201.
- Fluckiger, J., Vigil, Y. T. y., Pasco, R., & Danielson, K. (2010). Formative feedback: Involving students as partners in assessment to enhance learning. *College Teaching*, 58(4), 136–140.
- Fui, C. S., & Lian, L. H. (2018). The effect of computerized feedback on students' misconceptions in algebraic expression. *Pertanika Journal of Social Sciences and Humanities*, 26(3), 1387–1403.
- Goodman, D. P., & Hambleton, R. K. (2004). Student test score reports and interpretive guides: Review of current practices and suggestions for future research. *Applied Measurement in Education*, 17(2), 145–220.

- Gotch, C. M., & Roduta Roberts, M. (2018). Score reporting. In B. B. Frey (Ed.), *The SAGE encyclopedia of educational research, measurement, and evaluation*. Thousand Oaks, CA: Sage.
- Gutierrez, F., & Atkinson, J. (2011). Adaptive feedback selection for intelligent tutoring systems. *Expert Systems with Applications*, 38(5), 6146–6152.
- Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. New York, NY: Taylor & Francis.
- Hambleton, R. K., & Zenisky, A. L. (2013). Reporting test scores in more meaningful ways: A research-based approach to score report design. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J. I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology, Vol. 3: Testing and assessment in school psychology and education* (pp. 479–494). Washington, DC: American Psychological Association.
- Harks, B., Rakoczy, K., Hattie, J., Besser, M., & Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: the role of feedback's perceived usefulness. *Educational Psychology*, 34(3), 269–290.
- Hattie, J. (2009, April). *Visibly learning from reports: The validity of score reports*. Paper presented at the meeting of the National Council on Measurement in Education, San Diego, CA.
- Hattie, J., & Gan, M. (2011). Instruction based on feedback. In P. Alexander & R. E. Mayer (Eds.), *Handbook of research on learning and instruction* (pp. 249–271). New York, NY: Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hatziapostolou, T., & Paraskakis, I. (2010). Enhancing the impact of formative feedback on student learning through an online feedback system. *Electronic Journal of E-Learning*, 8(2), 111–122.
- Hepplestone, S., Holden, G., Irwin, B., Parkin, H. J., & Thorpe, L. (2011). Using technology to encourage student engagement with feedback: a literature review. *ALT-J: Research in Learning Technology*, 19(2), 117–127.
- Jonsson, A. (2012). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education*, 14(1), 63–76.
- Jug, R., Jiang, X., & Bean, S. M. (2019). Giving and receiving effective feedback: A review article and how-to guide. *Archives of Pathology & Laboratory Medicine*, 143(2), 244–250.
- Kulkarni, C. E., Bernstein, M. S., & Klemmer, S. R. (2015). *PeerStudio: Rapid peer feedback emphasizes revision and improves performance*. Vancouver, Canada: ACM.
- Kyllonen, P. (2009). New constructs, methods, and directions for computer-based assessment. In F. Sheuermann & J. Björnsson (Eds.), *The transition to computer-based assessment. New approaches to skills assessment and implications for large-scale testing* (pp. 151-156). Luxemburg: Office for Official Publications of the European Communities.
- Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, 23(2), 157–172.
- Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 75–106). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lilley, M. & Barker, T. (2007). Students' perceived usefulness of formative feedback for a computer-adapted test. *The Electronic Journal of e-Learning*, 5(1), 31–38.
- Lipnevich, A., & Smith, J. (2008). *Response to assessment feedback: The effect of grades, praise, and source of information* (ETS RR-08-30). Princeton, NJ: ETS.
- Lopez, L. (2009). *Effects of delayed and immediate feedback in the computer-based testing environment* [Electronic version]. Doctoral dissertation, Department of Curriculum, Instructional, and Media Technology, Indiana State University.
- Lowry, R. (2005). Computer aided self assessment – an effective tool. *Chemistry Education Research and Practice*, 6(4), 198–203.
- Maier, U., Wolf, N., & Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers & Education*, 95, 85–98.

- McCarthy, J. (2017). Enhancing feedback in higher education: Students' attitudes towards online and in-class formative assessment feedback models. *Active Learning in Higher Education*, 18(2), 127–141.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: American Council on Education.
- Moridis, C. N., & Economides, A. A. (2012). Applause as an achievement-based reward during a computerized self-assessment test. *British Journal of Educational Technology*, 43(3), 489–504.
- Moss, P. A. (1995). Themes and variations in validity theory. *Educational Measurement: Issues and Practice*, 14, 5–13.
- Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. van Merriënboer & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (pp. 125–143). New York, NY: Lawrence Erlbaum Associates.
- Nicol, D. (2007). E-assessment by design: Using multiple-choice tests to good effect. *Journal of Further and Higher Education*, 31(1), 53–64.
- Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–216.
- Pendleton, D., Schofield, T., Tate, P., & Havelock, P. (2003). *The new consultation: Developing doctor–patient communication* (2nd ed). Oxford, UK: Oxford University Press.
- Perie, M., & Huff, K. (2015). *Determining content and cognitive demand for achievement tests*. In S. Lane, M. R. Raymond, & T. Haladyna (Eds.), *Handbook of test development* (2nd ed., pp. 119–143), Routledge, UK: London.
- Prince, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: all that effort, but what is the effect? *Assessment & Evaluation in Higher Education*, 35(3), 277–289.
- Quinton, S., & Smallbone, T. (2010). Feeding forward: Using feedback to promote student reflection and learning: A teaching model. *Innovations in Education and Teaching International*, 47, 125–135.
- Race, P. (2001). *Using feedback to help students to learn*. The Higher Education Academy. Retrieved from [https://phil-race.co.uk/wp-content/uploads/Using\\_feedback.pdf](https://phil-race.co.uk/wp-content/uploads/Using_feedback.pdf)
- Race, P. (2006) *The lecturer's toolkit: A practical guide to assessment, learning and teaching*, (3rd ed.). London: Routledge.
- Schartel, S. A. (2012). Giving feedback—An integral part of education. *Best Practice & Research Clinical Anaesthesiology*, 26(1), 77–87.
- Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them*. New York, NY: W.W. Norton & Company.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189.
- Singh, U. G., & Villiers, M. R. (2017). An evaluation framework and instrument for evaluating e-assessment tools. *International Review of Research in Open and Distributed Learning*, 18(6), 165–185.
- Sinharay, S. (2010). How often do subscores have added value? Results from operational and simulated data. *Journal of Educational Measurement*, 47(2), 150–174.
- Slater, S., Livingston, S. A., & Silver, M. (2019). Score reports for large-scale testing programs. Managing the design process. In D. Zapata-Rivera (Ed.), *Score reporting research and applications* (pp. 91–106). New York, NY: Routledge.
- Suen, H. (2014). Peer assessment for massive open online courses (MOOCs). *The International Review of Research in Open and Distributed Learning*, 15(3), 312–327.
- Thurlings, M., Vermeulen, M., Bastiaens, T., & Stijnen, S. (2013). Understanding feedback: A learning theory perspective. *Educational Research Review*, 9, 1–15.
- Timmers, C. F. (2013). *Computer-based formative assessment: Variables influencing feedback behaviour* (Unpublished doctoral dissertation). University of Twente, Enschede, the Netherlands.

- Timmers, C. F., & Veldkamp, B. P. (2011). Attention paid to feedback provided by a computer-based assessment for learning on information literacy. *Computers & Education, 56*, 923–930.
- Van der Kleij, F. M., Eggen, T. J. H. M., Timmers, C. F., & Veldkamp, B. P. (2012). Effects of feedback in a computer-based assessment for learning. *Computers & Education, 58*, 263–272.
- Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes a meta-analysis. *Review of Educational Research, 85*(4), 475–511.
- Walker, D. J., Topping, K., & Rodrigues, S. (2008). Student reflections on formative e-assessment: Expectations and perceptions. *Learning, Media and Technology, 33*(3), 221–234.
- Weedon, E. (2000). Do you read this the way I read this? *British Journal of Educational Technology, 31*, 185–197.
- Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco, CA: Jossey-Bass.
- Wu, H. M., Kuo, B. C., & Wang, S. C. (2017). Computerized dynamic adaptive tests with immediately individualized feedback for primary school mathematics learning. *Educational Technology & Society, 20*(1), 61–72.
- Wuensch, K. L., Aziz, S., Ozan, E., Kishore, M., & Tabrizi, M. H. N. (2008). Pedagogical characteristics of online and face-to-face classes. *International Journal on E-Learning, 7*(3), 523–532.
- Yuan, J., & Kim, C. (2015). Effective feedback design using free technologies. *Journal of Educational Computing Research, 52*(3), 408–434.
- Zapata-Rivera, D., & Katz, I. R. (2014). Keeping your audience in mind: Applying audience analysis to the design of interactive score reports. *Assessment in Education: Principles, Policy & Practice, 21*, 442–463.
- Zenisky, A. L., & Hambleton, R. K. (2012). Developing test score reports that work: The process and best practices for effective communication. *Educational Measurement: Issues and Practice, 31*, 21–26.
- Zenisky, A. L., & Hambleton, R. K. (2016). A model and good practices for score reporting. In S. Lane, M. R. Raymond, & T. M. Haladyna (Eds.) *Handbook of test development* (2nd ed., pp. 585–602). New York, NY: Routledge.