3-17-2022

Evaluating Feedback Post-Computer-Based Assessment in Health Professions Education: A Systematic Review

Bethany Huebner
Bellarmine University, bhuebner@bellarmine.edu

Evaluating feedback post-computer-based assessment in health professions education: a systematic review

Bethany Huebner[1,2], Barbara Jackson[2], Megan Danzl[2], Jason Pitt[1]

[1]Department of Physical Therapy, University of Evansville, Evansville, Indiana, USA, [2]PhD in Health Professions Education Program, Bellarmine University, Louisville, Kentucky, USA

Correspondence: Bethany Huebner, University of Evansville, Department of Physical Therapy, 515 Bob Jones Way, Evansville, IN, 47630, USA. Tel: 1-812-488-4062; email: bh212@evansville.edu

Evaluating feedback post-computer-based assessment in health professions education: a systematic review

Abstract

Background: Computer-based assessments are common in health professions education and offer robust feedback options. The style of feedback that is best for student learning is unclear.

Aim: To systematically review feedback post-computer-based assessment literature to identify key feedback strategies to optimize student learning and retention.

Methods: A search of electronic library databases, a supplemental Internet source, and reference lists were completed. Inclusion criteria were any English-language sources that used feedback post-computer-based assessment. Data were analyzed qualitatively and summarized descriptively.

Results: There were 134 records identified for initial relevancy through screening by title and abstract. The full texts of 41 records were retrieved and assessed further for relevance. A total of 23 records were analyzed. Three major themes were identified: Types of feedback, the timing of feedback, and student utilization.

Conclusion: Feedback post-computer-based assessments are an essential part of student learning. The type and timing of feedback delivery should be considered, and student engagement with feedback.

Keywords: computer-based exam, feedback, elaborated feedback

Introduction

Computer-based testing streamlined the testing process for higher education programs by creating an efficient, robust, and secure testing process for student assessment. Computer-based testing software companies market the rich data produced for analysis post examination, the improved timing of results, the lengths of exam security, and the robust feedback students may receive (Zheng and Bender 2019). Licensure examinations, like the National Physical Therapy Examination, adopted computer-based testing to replace paper examinations and improve examination security. With this update, exam takers are provided with more efficient results and a standardized licensure process (A History of the Physical Therapy Examination | FSBPT 2019). After reviewing the evidence, researchers suggest that faculty utilize computer-based testing for ease of use and improved feedback, and students prefer it for similar reasons (Pawasauskas et al. 2014; Wadley et al. 2014; Bloom et al. 2018). However, few systematic reviews examine which computer-based examination style of feedback most benefits students' learning and retention.

Feedback is the glue that holds the learning process together. The surge of published articles over the past 50 years on feedback supports its use as one of the most potent influences on the learning process (Shute 2008; Pawasauskas et al. 2014; Wadley et al. 2014; Bloom et al. 2018). If formative, feedback provides information for the student to modify their behavior or thought process to improve learning (Shute 2008; Bennett 2011; Carless and Boud 2018). Knowledge of results (KR), which provides feedback on the correctness of the task, is most associated with the computer-based post-examination feedback (Shute 2008). In addition to students receiving KR, some computer-based testing software provides categorical feedback, enhanced feedback, and even a secure examination review (Hattie and Timperley 2007; Simpson

LP and Justice 2016; Simpson L et al. 2018; Zheng and Bender 2019). In health professions programs, KR is most often provided post formative assessments. These assessments provide steppingstones as the student learns new knowledge and skills and applies and synthesizes information to patient cases (Maier et al. 2016).

This variety of options without evidence of best practices raises challenges for a faculty member to determine an optimal approach to provide feedback to students post-computer-based exam taking. We, therefore, systematically studied the health professions education literature to identify best practices for giving and utilizing feedback post-computer-based assessments.

Theoretical Framework

In a comprehensive literature review related to feedback, Hattie and Timperley (2007) layout a framework of feedback that incorporates the student, instructor, and the various levels in which feedback is provided (Figure 1). Hattie and Timperley (2007) report that the primary purpose of feedback is to "reduce discrepancies between current understandings and performance and a goal." (pg 86). Their framework requires instructors to think about feedback from answering three specific questions, "Where am I going? (What are the goals?), How am I going? (What progress is being made toward the goal?), and Where to next? (What activities need to be undertaken to make better progress?)."

Simply answering those questions does not fully provide the depth in feedback needed; however, within answering these questions, the instructor is also challenged to consider which level the feedback provided operates. In their model, Hattie and Timperley (2007) outline four levels in their framework: task level, process level, self-regulation level, and self-level. The task-level focuses on the actual task or skill. The process level focuses on the process utilized to complete the task. The self-regulation level focuses on how the learner self-reflects and assesses where they are in the learning process. The self-level focuses on the learner and their affect. Hattie and Timperley (2007) theorize that designing feedback to answer the three main questions and operating at all four levels should provide effective feedback for learning.

This framework guides this study regarding feedback and computer-based exam-taking by utilizing the framework's concepts to search for literature that identifies the most effective feedback form for students' learning. The traditional approach to feedback post-examination includes KR primarily, but that type of feedback only operates at the task level. According to Hattie and Timperley's (2007) theory, only providing KR would not be an effective form of feedback because it does not answer all three questions or operate at all four levels. Other potential forms of feedback post-computer-based assessments like categorical feedback, secure examination review, or enhanced feedback may meet the model's theoretical needs and be considered more preferred. A systematic review of the current literature with the lens of this model may provide educators with a better process for delivering effective feedback to the learners.

Methods

A comprehensive inquiry of the scholarly sources was completed to identify evidence-based best practices for feedback delivery post-computer-based assessments in graduate health profession education programs.

Inclusion/Exclusion Criteria

Inclusion criteria for this systematic review included: peer-reviewed published sources, English language sources, sources utilizing a study design that had feedback assessment post-computer-based exams, sources using feedback post formative assessments, sources published since 2007, and sources that contained a population of health professions graduate students. All studies that met our inclusion criteria were included in the review. No additional exclusion criteria were placed on study design or health profession field.

Search Strategy

An electronic database search strategy was utilized to identify relevant studies. The primary electronic database used was EBSCOhost Academic Libraries accessed through W.L. Lyons Brown Bellarmine University Library. All databases in EBSCOhost were included. Google Scholar database was also utilized for a secondary search strategy. A third strategy was reviewing references in found studies, and those titles were accessed with a hand search. Searches were conducted using a combination of search terms: "feedback," "computer-based assessment," "student learning," "computer-assisted testing," "graduate students," "health professions," "computer-based exams," "formative assessment," "elaborated feedback," and "higher education." The search was completed in October 2021.

Selection Methods

The titles and abstracts were examined for relevance to the research question. Records were considered potentially relevant if they included computer-based assessment and feedback-related terms. Where the information was not available to determine if an article met the inclusion criteria after reviewing of full-text, it was rejected. All potentially relevant records were retrieved in full-text and screened for final inclusion.

Data Analysis

Due to the various studies included, there was no universal scoring method available for the authors; therefore, a qualitative approach was utilized. Data extraction was based on Decuir-Gunby et al. Field (2010) developed techniques. The process for reviewing articles to develop codes focuses on reading, note-taking, re-reading, and more note-taking. This process can be broken down into open and axial coding fields (DeCuir-Gunby et al. 2010). Open coding is the first step, in the beginning, to formulate data from the articles. This process allows the researcher to explore ideas and highlight critical areas. This process does not entirely focus on correlating data sets, as this is just the opening step to developing themes. Axial coding begins the process of narrowing and connecting amongst data sets. This process is a higher level of coding and helps the researcher establish connections between codes (DeCuir-Gunby et al. 2010).

Results

Upon initial screening, 134 studies were identified via electronic databases and hand searching. After reviewing the title and abstracts, 24 duplicate and 69 irrelevant records were removed. The full-text reports of 41 articles were retrieved and assessed for relevance. Of the 40 relevant records, 23 met the inclusion criteria and formed the basis of our analysis. See Figure 2 for PRISMA results. Three main themes were identified. The first theme was the type of feedback given post-computer-based testing: knowledge of results (KR), knowledge of correct results (KCR), or elaborated feedback (EF). The second theme that emerged was the timing of feedback. The third theme that surfaced was student utilization of feedback.

Discussion

Type of Feedback

With the ability to produce robust assessment results, computer-based assessment platforms can provide a variety of feedback options (Butler A. C. and Roediger 2008; van der Kleij et al. 2012; Malau-Aduli et al. 2013; Van der Kleij et al. 2015; Attali and van der Kleij 2017; Petrović et al. 2017; Levant et al. 2018; Zheng and Bender 2019). Although there can be varying options, Shute (2008) identified three main categories of feedback associated with computer-based assessment: knowledge of results (KR), knowledge of correct response (KCR), and elaborated feedback (EF). KR informs the student whether the answer they chose was correct or incorrect; it does not provide the correct response or any additional information (Van der Kleij et al. 2015). KCR goes one step further and informs the student of the correct answer. Both KR and KCR have a corrective focus. On the other hand, EF is not as corrective as KR or KCR but more informative. EF is best known as a new form of instruction, and it often coincides with KR or KCR. EF is defined variably but is known for the ability to take feedback one step further into further instruction with either worked out solutions, answer rationale or categories for further development (Van der Kleij et al. 2015).

In a meta-analysis, Van der Kleij et al. (2015) identified that providing students with EF led to better learning outcomes than KR or KCR. They also determined that EF performed better with higher-order learning outcomes than lower-order and simple feedback (Van der Kleij et al. 2015). In a follow-up study, Attali and van der Kleij (2017) further identified EF as beneficial primarily in cases where students had no prior knowledge of a topic. Other researchers have also identified a link between EF and positive student learning outcomes post-computer-based assessments (Butler A. C. and Roediger 2008; Maier et al. 2016; Petrović et al. 2017; Levant et al. 2018; Zheng and Bender 2019). Levant et al. (2018) utilized EF by including question rationales with KCR for medical students and was able to identify a significant improvement in re-test performance when re-testing on related questions, not exact repeat questions. In this study, the researchers only exposed the students to EF if they missed the question, which differed from the Attali and van der Kelij (2017) study that gave EF for all questions.

An analysis of these articles indicates EF is the ideal type of feedback to provide students post-computer-based assessments to improve learning outcomes on re-test (Van der Kleij et al.

2015; Maier et al. 2016; Attali and van der Kleij 2017; Levant et al. 2018). Although that seems definitive, EF is quite varied in its definition, which is problematic for educators who want to deliver consistent and beneficial feedback to students. Providing students with KR or KCR may provide a better learning outcome as compared to no feedback, but what else needs to be included in to complete EF is yet to be fully defined (Butler A. C. and Roediger 2008; Van der Kleij et al. 2015; Maier et al. 2016; Attali and van der Kleij 2017; Levant et al. 2018; Zheng and Bender 2019).

Given the various forms of EF, we sought guidance from researchers to identify components of effective feedback. In their model of feedback, Hattie and Timperley (2007) theorize that effective feedback must answer three questions, "Where am I going? (What are the goals?), How am I going? (What progress is being made toward the goal?), and Where to next? (What activities need to be undertaken to make better progress?)." In addition to answering those specific questions, educators need to consider what level of learning to direct the feedback to: task, process, self-regulation, or self (Hattie and Timperley 2007). In their meta-analysis, Van der Kleij et al. (2015) identified EF that included feedback oriented to the task and process level to demonstrate moderate effect sizes (ES'=0.50, k=41), and feedback involving the task and self-regulation levels showed the highest effect size (ES'=1.05, k=4). These feedback types were compared to a similar group of students who only received KR, and there was a significant difference between those groups. More specific studies investigating feedback that answer Hattie and Timperley's (2007) three questions at all four levels would benefit educators as they make decisions on feedback delivery.

Timing of Feedback

Upon reviewing the relevant articles, the theme of the timing of feedback emerged (Gibbs and Simpson 2004; Butler Andrew C. et al. 2007; Hattie and Timperley 2007; Shute 2008; Attali and van der Kleij 2017; Levant et al. 2018). The timing of feedback is an essential aspect of the learning process as students must receive feedback in the window before moving on to new content (Gibbs and Simpson 2004). Just like EF definitions are variable, defining immediate and delayed feedback is inconsistent in the literature. Van der Kleij et al. (2012) attempted to create a working definition related to computer-based assessments. Immediate feedback occurs directly after the learner responds to a question, and delayed feedback occurs any time after all items are completed. Delayed feedback encompasses as short as seconds after turning in the entire test to weeks or months after completing the assessment. These variations in delayed feedback make it difficult to fully assess timing as a factor in feedback to improve learning outcomes.

Inconsistent effects of feedback timing on learning outcomes are reported. Van der Kleij et al. (2015) promote immediate feedback as superior to delayed feedback; however, other researchers found delayed feedback superior to immediate feedback (Butler Andrew C. et al. 2007; Levant et al. 2018). Levant et al. (2018) found that providing a secure exam review where students could see the entire question with answers and rationale (EF) for 30 minutes immediately after completing the assessment supported the significant improvement of re-test scores. Butler et al. (2007) also identified delayed EF by 1-day post-assessment demonstrated a significant improvement in re-test scores instead of immediate feedback. With a point of contention for and against, Corbett and Anderson (2001) proposed a theory that feedback timing may need to be matched to the type of content being learned. Their research identified delayed feedback as better for promoting transfer of learning, whereas immediate feedback is better for learning tasks or procedural type skills (Corbett and Anderson 2001). These inconsistent results are not abnormal for this variable with feedback as the effects of immediate versus delayed feedback on learning was a point of contention for decades (Shute 2008).

With some evidence demonstrating significant effects with immediate feedback and some for delayed feedback, educators must use the best judgment to identify the most appropriate timing for students. It is important to note that even though forms of delayed feedback demonstrated improvement in re-test scores, the longest delay studied was only one day. Gibbs and Simpson (2004) urge educators to provide timely feedback while it still matters to the student. Often after assessments, the curriculum moves forward, and if students do not receive feedback promptly, the students may not attune to the feedback even once received. If a student does not attune to feedback, the learning loop cannot be closed, nor can the student benefit from the feedback (Gibbs and Simpson 2004).

Student Utilization of Feedback

The final theme is student utilization of feedback as a variable that improves learning outcomes post-computer-based assessments. Regardless of feedback type or timing, having a student actively involved in feedback reception is essential for correcting errors and demonstrating learning (Nicol and Macfarlane-Dick 2006). Instructors primarily give feedback, but the learner must be open to receiving it. In addition, learners need to understand the feedback and apply it to learning and correcting errors (Nicol and Macfarlane-Dick 2006; Boud and Molloy 2013; Malau-Aduli et al. 2013; Maier et al. 2016; Lam 2017; O'Donovan 2017). Maier et al. (2016) assessed various factors influencing student feedback utilization through a survey. Two areas they identified as predictive factors in lower utilization of feedback post-assessment were individual differences in motivation and cognition and generally low interest in a subject (Maier et al. 2016).

When looking at feedback post-computer-based assessment mainly, Karay et al. (2012) found that students were more engaged in utilizing feedback when the feedback was more informative. Direct and detailed feedback (EF) was better received and engaged with by their sample, increasing the student acceptance of taking assessments on a computer (Karay et al. 2012). Zheng and Bender (2019) further identified EF post-computer-based assessment to significantly predict if students accepted the utilization of computer-based assessment protocols over paper-based assessment protocols. In focus groups, students identified the timeliness of feedback post-assessment and clarity of feedback as primary reasons for interacting with the feedback and utilizing it for future assessments (Zheng and Bender 2019). Shute (2008) states that for feedback to effectively engage students with the process, it requires three things: 1. Motivation, 2. Opportunity, and 3. Means (Shute 2008). Student motivation is necessary to hear and internalize feedback, students need the opportunity to review feedback, and educators need to provide quality feedback. O'Donovan et al. (2017) found that most students appreciate feedback and often crave it; however, they do not always find it helpful in their learning and growth. Many students fail even to read the feedback provided (O'Donovan 2017) or perceive the feedback as too lengthy or riddled with complex ideas to utilize properly (Shute 2008). Without carefully incorporating the three requirements of effective feedback, much of the message to improve or change knowledge or behavior is lost as the students fail to engage. O'Donovan et al. (2017) also identified that many students, although motivated to receive feedback, do not know how to interpret it or utilize it for their growth. Literature on student engagement with feedback is growing. Still, it continually points to students needing training on interpreting feedback (O'Donovan 2017).

Conclusion

Feedback is an essential component of the learning process that requires all stakeholders to participate in various ways. Both educators and students have roles to play in the process of giving and receiving feedback. When providing feedback, educators may use these questions as a good guide to provide effective feedback: 1. Where am I going? 2. How am I going? and 3. Where to next? (Hattie and Timperley 2007). Further, they will want to consider which level (task, process, self-regulation, or self) the answers to those questions are operating. Utilizing this model will cause educators to use EF as a post-assessment feedback tool.

Educators may find the best learning outcomes when using EF over KR or KCR when utilizing computer-based assessments. Educators can create clear and engaging EF post-computer-based assessments following guidelines in the literature and instruct students on how best to use the feedback once it is received. Regarding timing, it is difficult to provide a recommendation on when feedback should be delivered, either immediate or delayed, as there are mixed outcomes. Delayed feedback, given any time within one day after an assessment is completed, may have a slight edge of demonstrating improved learning outcomes than immediate feedback. However, the best recommendation is to provide feedback promptly so students can apply it before moving on to new material (Van der Kleij et al. 2015). Finally, educators may consider more detailed and direct feedback to improve student utilization, so students can clearly understand how to enhance their learning gaps. Students also need to engage in the feedback and apply it to their learning. In the future, researchers may consider if training on feedback styles improves student utilization over time and if a more specific type of feedback besides the broad definition of EF is more conducive to student outcomes.

References

Attali Y, van der Kleij F. 2017. Effects of feedback elaboration and feedback timing during computer-based practice in mathematics problem solving. Computers & Education. 110:154-169.

Bennett RE. 2011. Formative assessment: a critical review. Assessment in Education: Principles, Policy & Practice. 18(1):5-25.

Bloom TJ, Rich WD, Olson SM, Adams ML. 2018. Perceptions and performance using computer-based testing: One institution's experience. Curr Pharm Teach Learn. 10(2):235-242.

Boud D, Molloy E. 2013. Rethinking models of feedback for learning: the challenge of design. Assessment & Evaluation in Higher Education. 38(6):698-712.

Butler AC, Karpicke JD, Roediger HL. 2007. The effect of type and timing of feedback on learning from multiple-choice tests. Journal of Experimental Psychology: Applied. 13(4):273-281.

Butler AC, Roediger HL, 3rd. 2008. Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. Mem Cognit. 36(3):604-616.

Carless D, Boud D. 2018. The development of student feedback literacy: enabling uptake of feedback. Assessment & Evaluation in Higher Education. 43(8):1315-1325.

Corbett AT, Anderson JR. 2001. Locus of feedback control in computer-based tutoring: impact on learning rate, achievement and attitudes. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Seattle, Washington, USA.

DeCuir-Gunby JT, Marshall PL, McCulloch AW. 2010. Developing and using a codebook for the analysis of interview data: An example from a professional development research project. Field Methods. 23(2):136-155.

Gibbs G, Simpson C. 2004. Does your assessment support your students' learning? Learning and Teaching in Higher Education. 1(1):1-31.

Hattie J, Timperley H. 2007. The power of feedback. Review of Educational Research. 77(1):81-112.

A History of the Physical Therapy Examination | FSBPT. 2019. [accessed 2019 September 3]. https://www.fsbpt.org/Free-Resources/NPTE-Articles-Old/A-History-of-the-Physical-Therapy-Examination.

Karay Y, Schauber SK, Stosch C, Schuettpelz-Brauns K. 2012. Can computer-based assessment enhance the acceptance of formative multiple choice exams? A utility analysis. Med Teach. 34(4):292-296.

Lam R. 2017. Enacting feedback utilization from a task-specific perspective. The Curriculum Journal. 28(2):266-282.

Levant B, Zuckert W, Paolo A. 2018. Post-exam feedback with question rationales improves re-test performance of medical students on a multiple-choice exam. Adv Health Sci Educ Theory Pract. 23(5):995-1003.

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.

Malau-Aduli BS, Assenheimer D, Choi-Lundberg D, Zimitat C. 2013. Using computer-based technology to improve feedback to staff and students on MCQ assessments. Innovations in Education and Teaching International. 51(5):510-522.

Nicol DJ, 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):199-218.

O'Donovan B. 2017. How student beliefs about knowledge and knowing influence their satisfaction with assessment and feedback. Higher Education: The International Journal of Higher Education Research. 74(4):617-633.

Pawasauskas J, Matson KL, Youssef R. 2014. Transitioning to computer-based testing. Currents in Pharmacy Teaching and Learning. 6(2):289-297.

Petrović J, Pale P, Jeren B. 2017. Online formative assessments in a digital signal processing course: Effects of feedback type and content difficulty on students learning achievements. Education and Information Technologies. 22(6):3047-3061.

Shute VJ. 2008. Focus on formative feedback. Review of Educational Research. 78(1):153-189.

Simpson L, Jones M, Taylor S. 2018. Evaluating the Use of Automated Feedback on Student Motivation. INNOVATIONS in pharmacy. 9(2).

Simpson LP, Justice J. 2016. Perception of examsoft feedback reports as autonomy-support for learners.

van der Kleij FM, Eggen TJHM, Timmers CF, Veldkamp BP. 2012. Effects of Feedback in a Computer-Based Assessment for Learning. Computers & Education. 58(1):263-272.

Van der Kleij FM, Feskens RCW, Eggen TJHM. 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.

Wadley M, Weaver SB, Curry C, Carthon C. 2014. Pharmacy students' perceptions of ExamSoft® as the primary assessment tool in an integrated therapeutics course. Currents in Pharmacy Teaching and Learning. 6(6):815-821.

Zheng M, Bender D. 2019. Evaluating outcomes of computer-based classroom testing: Student acceptance and impact on learning and exam performance. Med Teach. 41(1):75-82.

Figure 1: Feedback theory adapted from Hattie and Timperley (2007)