Guidelines for Online Assessment in Emergency Remote Teaching during the COVID-19 Pandemic

Ahmad Fuad Abdul Rahim

Medical Education Department, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia

To cite this article: Rahim AFA. Guidelines for online assessment in emergency remote teaching during the COVID-19 pandemic. Education in Medical Journal. 2020;12(2):59–68. https://doi.org/10.21315/eimj2020.12.2.6

To link to this article: https://doi.org/10.21315/eimj2020.12.2.6

ABSTRACT

The COVID-19 pandemic has precipitated a shift to emergency remote teaching (ERT) for educational institutions worldwide, including medical schools. To help faculty design online assessments, guidelines available in the internet were compiled and nine common themes were identified. These nine guidelines for online assessment were discussed in the light of ERT as well as relevance to health professions education. The guidelines include: Evaluate prerequisites for implementing online assessment; ensure alignment of assessment activities with stated learning objectives; address the diversity of students’ situations; maintain a good balance of formative and summative assessments; stimulate student learning with online assessment; consider format; scheduling and timing of tests; establish clear communication to students regarding assessment matters; ensure high-quality feedback; and address assessment validity threats.

Keywords: Assessment, Online teaching, Emergency remote teaching, COVID-19, Pandemic

INTRODUCTION

The current COVID-19 pandemic impacts on all aspects of human life, including education. Reports stated that from 290.5 million (1) to up to 1.725 billion students (2) are affected by closures of educational institutions, involving just over 100 countries (3) spanning three continents (1, 4–7). Most medical schools throughout the world have also closed (8).

Educational institutions worldwide have responded to closures in many ways, but by far the almost universal reaction is shifting to online teaching and learning (OT&L) (4–5). Although the scope, intensity and methods vary between countries and institutions, common features include the usage of fully remote teaching methods which are reliable and easy to set up, replacing previous face-to-face or other combinations of teaching methods. Additionally, there is a stated or unstated understanding that the measures are temporary and educational institutions will revert to previous methods after the current crisis is over. The term emergency remote teaching (ERT) was proposed to distinguish it from established, well-designed and long-term online teaching (9). This sudden change to ERT may be stressful to faculty...
(9) who are not familiar with OT&L concepts and methods, especially senior faculty members and busy clinicians.

One of the areas that faculty needs to contend with is the design of online assessment. Guidelines for designing online assessment are available at various sites on the internet but are variable in content. These guidelines were compiled, and nine common themes were derived. In this article, these nine guidelines were then discussed in the light of ERT. Whenever relevant, situations specific to assessment in medical training are discussed. It is hoped that this will be a useful guide for faculty when designing online assessment.

GUIDELINE 1: EVALUATE PREREQUISITES FOR IMPLEMENTING ONLINE ASSESSMENT

Prerequisites are important if an online assessment is to be successful. Two categories of prerequisites have been identified; institutional and educator readiness (10).

Institutional readiness encompasses institutional policies, resources and practices. In the light of the pandemic, many institutions place their hopes of survival on online teaching, therefore policy is often very supportive. More problematic, however, is the issue of the infrastructure needed for online teaching, especially internet coverage. Globally only about 60% of the population has internet access (11), even in developed countries such as Britain (about 90%) and Japan (about 80%) (3). In the United States, 21.3 million Americans still lack adequate broadband access in rural areas (12). At the institutional level, institutions with an online learning management system (LMS) in place have a head start in terms of information technology (IT) support plus faculty and student familiarity. However, the massive and sudden shift to online teaching can overwhelm the available IT support team (9). For those without such assets in place, a huge array of options is available; some of which are free (10).

Educator readiness is equally important. Excellent infrastructure and available software and hardware are useless without convincing faculty about the necessity and benefits of the shift (10). Also, they need to be trained in sound assessment practices and the use of available online assessment applications.

For both categories, the role of a responsive institutional leadership is essential (13–14) to determine baseline infrastructural and human resource preparedness, decide on and clearly communicate a strategy for information and communications technology (ICT) integration, organise faculty development activities in ICT integration and online assessment and ensuring adequate technical support for educators when using online assessment tools.

GUIDELINE 2: ENSURE ALIGNMENT OF ASSESSMENT ACTIVITIES WITH STATED LEARNING OBJECTIVES

A fundamental principle of assessment is that it must be designed to assess stated learning objectives (15–24). Cognitive learning objectives are often stated using Bloom’s Taxonomy, which was revised in 2001 to address 21st-century skills (10). Authors have even created a Bloom’s Digital Taxonomy (25). A detailed discussion of the taxonomy and suitable assessment tools for each level is out of the scope of this guide, but readers are referred to excellent resources such as Westhuizen’s ‘Guidelines for Online Assessment for Educators’ (10).

The assessment of the affective and psychomotor domain in the ERT situation is difficult, especially in medical education assessment for outcomes such as clinical skills, professionalism, empathy and
teamwork. Valid assessment of these outcomes, commonly using methods such as objective structured clinical examinations (OSCEs) and short- or long-cases involves face-to-face observation (26–27) which is not possible in the scenario of school closures and social-distancing orders. Many international medical training boards, colleges and councils have cancelled or postponed clinical examinations in the light of this restriction (28–33) aiming instead for a more favourable situation in the near future. However, there is still room for creativity in this regard; there is a cognitive component in clinical skills (26) and institutions might decide to focus on this in the current situation. History-taking and patient counselling skills are assessable using video-conferencing applications, as demonstrated by Qatar University’s College of Pharmacy (34). Performance tests, however, are still difficult to assess without introducing serious threats to validity (27).

GUIDELINE 3: ADDRESS THE DIVERSITY OF STUDENTS’ SITUATIONS

In ERT, students access teaching and learning from diverse situations in terms of their geographical locations and other factors. Designers of assessments should be aware of this diversity so as not to discriminate students with poor support for OT&L.

At this point, it is useful to differentiate synchronous and asynchronous OT&L methods. Synchronous OT&L methods allow learning to occur much like a class situation (35). Educators and students can see and interact with each other at a specified time using methods such as video conferencing or live chatting. Its advantages include being highly engaging and dynamic and if done frequently it allows deeper relationships between educators and students. Assessment-wise, oral assessment methods like vivas are possible. On the flip side, synchronous methods often require good internet facilities and good time management. On that note, even multiple-choice questions, if done simultaneously at a specified time and date should be considered as synchronous and therefore the same considerations about discriminating students with poor ICT infrastructures apply.

Asynchronous OT&L methods are not real-time based (35). Learning materials are uploaded, and students access them whenever convenient or whenever the internet coverage allows. Methods include uploading videos, lecture notes and social-media-style interactions. Its advantages of flexibility and allowing student self-pacing are offset by the absence of real-time and face-to-face feedback. Asynchronous assessment methods include assignments and projects.

Due to the requirement for good internet quality in synchronous OT&L, asynchronous methods are given priority when students have varied levels of ICT infrastructure quality. If synchronous assessments are deemed necessary, options for students without access to good internet facilities should be made available, such as designating centres with good internet facilities that students can go to for their examinations.

Diversity is not limited to ICT infrastructure. Assessment designers should also consider other variables such as socioeconomic status, second-language English speakers and students with special needs (10, 20, 24). Care also needs to be given to ensure that skills related to computer use, such as keyboard typing skills, and familiarity with certain applications will not give undue advantage to certain students over others (24). In ERT diversity of ICT infrastructure is the foremost concern; however, these other factors also need to be judiciously considered.
GUIDELINE 4: MAINTAIN A GOOD BALANCE OF FORMATIVE AND SUMMATIVE ASSESSMENTS

Assessment designers need to balance formative and summative assessments (10, 18). Formative assessments provide useful feedback to students regarding their achievement of stated learning objectives; they often occur during the course of study and are of low stakes (36). OT&L is very suitable for this purpose (37). Examples include weekly short quizzes. Summative assessments have the primary purpose of determining the achievement of students at or near the end of a course of study and are of high stakes (36). They include final professional or licensing examinations in medicine.

In the shift to ERT, students might feel disoriented. Frequent and diverse low-stakes formative assessments, such as quizzes help students familiarise themselves to online assessment methods and increases engagement which in turn increases learning (19, 20, 22, 38–41). It also helps them gain confidence (15) and prevents them from falling behind (20).

The need for stringent summative assessments is important in professional training such as in medicine. However, literature regarding summative assessments in OT&L is limited (37). Perhaps the biggest concern in this regard is examination security, which will be discussed below.

GUIDELINE 5: STIMULATE STUDENT LEARNING WITH ONLINE ASSESSMENT

Assessment methods affect how students learn (42), and online assessments must be designed to stimulate positive learning behaviours in students. This is apparent in the preceding discussion about formative assessment. In this aspect, other themes of online assessment-stimulated learning include encouraging collaboration (10, 17, 23, 43) by utilising group-based assessments. Another theme is the assessment of 21st-century skills which include learning and innovation, life and career, information media and technology (10). Interested readers are referred to The Partnership for 21st Century Learning (www.p21.org) for clarification of these skills.

Another theme which is perhaps more relevant in medical assessments is the emphasis on authentic assessment, in which assessment tasks are designed to reflect real-life, practical situations (10, 16, 23, 44). As discussed above, in summative assessments of the cognitive domain this means favouring the use of context-rich, real-life scenarios such as patient cases in multiple-choice questions and essays (20, 45). This is thought to better assess problem-solving and other higher-order thinking skills which in turn stimulate students to learn accordingly (42).

GUIDELINE 6: CONSIDER FORMAT, SCHEDULING AND TIMING OF TESTS

Formats or types of online assessments can initially be bewildering but as discussed above they can be grouped according to Bloom’s taxonomy levels, formative versus summative, or both (10, 46). Whenever relevant and feasible, using more than one format for a particular domain is recommended to measure widely and deeply (24). As with all assessments, quality-control procedures including blueprinting are required to ensure adequate coverage and sampling of all learning outcomes.

Online examinations should be scheduled carefully, considering the course timetable and other important events in the students’ calendar (15, 24). Both formative and summative examinations should have adequate time limits. For formative examinations, allowing multiple attempts improves student learning (19).
GUIDELINE 7: ESTABLISH CLEAR COMMUNICATION TO STUDENTS REGARDING ASSESSMENT MATTERS

Clear communication to students regarding assessment matters is considered as evidence supporting the validity of the assessment (47). It also helps students plan their learning and thus aid in developing self-management competencies (21). Besides, the shift from face-to-face teaching to ERT may be disruptive to students (40). They might be apprehensive of online assessments, especially for high-stakes examinations (48). Clear communication helps to alleviate these problems.

Areas of assessment that need to be clarified include how activities and assignments are related to course objectives and outcomes (40). The overall assessment specifications including the assessment schedule, types of assessment, the weightage of each method and the criteria for passing need to be sufficiently clarified (15, 16, 19, 21, 43). For assessments rated by faculty – such as assignments or projects – the rubrics need to be transparent (24, 40). For specific assessment methods, information such as when they will be opened and closed, duration of the examination, types of questions used, and the number of attempts allowed need to be stated. Feedback regarding examinations also needs to be clarified, such as how soon they will get their grades and how they can access it (40). Also, inform them of the procedure or the contact person in case of technical difficulties during examinations (19).

To increase clarity, some authors recommend explaining assessment matters using text, audio or video and providing examples whenever relevant (40) or making a FAQ section (19). It is also good to include them in decision-making discussions pertaining to assessment (43).

GUIDELINE 8: ENSURE HIGH-QUALITY FEEDBACK

Online assessment supports learning in the presence of timely, rich and constructive feedback (10, 20, 24, 49) and feedback makes the assessment effective (16).

Timely feedback allows students to correct themselves (40) and helps learning while promoting and maintaining student motivation (19). The content of the feedback includes information about the correctness of the students’ response in multiple-choice or true-false questions. This can be written into the LMS assessment system and is automatically given after students complete the examination. Additional explanations about the answers are recommended. Further information on how to improve performance is also useful (20) and can be in the form of frequent, low-stakes quizzes or on-line chats (40) or referrals to particular learning materials for further study. For optimal student learning, individualised student learning is preferable (19) and can be done efficiently via e-mail or instant-messaging apps. In ERT, it’s also important to acknowledge receipt of student assignments or projects (40).

GUIDELINE 9: ADDRESS ASSESSMENT VALIDITY THREATS

Validity is equally important in online assessments. Threats to validity include context under-representation (CU), which relate to issues of sampling and blueprinting, and construct-irrelevant variance (CIV), which relate to item quality, passing score determination and integrity of assessments, including issues of cheating and examination security (50).

Cheating is a concern especially in summative assessments in professional courses such as medicine. Research data on cheating in online assessments is
inconclusive. On the one hand, cheating seems to be a real cause for concern. Rowe (51) cited reports of approximately 70% of American high school seniors admitting to cheating. On the other hand, even if the prevalence is high, it had not been conclusively shown that cheating in online environments is significantly more prevalent than in conventional assessments (20, 52). Cheating in online assessment is not easier than other assessment situations (52); in fact, proctoring online examinations via human and software means have become big business (53) and cheating have actually become more difficult (54).

Strategies to overcome this threat begin by encouraging honesty among students (19). They need to be reminded about an academic integrity policy and made to agree to it (55). Clear communication regarding outcomes and expected performance also helps (49). The weight of final summative examinations can also be reduced and carefully redistributed to continuous assessments to promote an emphasis on positive learning behaviours (55). The use of open-book tests has been recommended (17, 56) although in ERT medical schools need to balance its advantages with the need to train their faculty to shift to a new assessment mindset. In cognitive assessment, multiple-choice questions with context-rich stems are favoured (20, 45) as well as open-ended questions incorporating case scenarios (51, 55–57). Increasing the use of short-notice, pop-quiz-style assessments is also suggested (55). Proctoring of online tests is currently possible using human proctors monitoring students online or by using software (53, 55) but purchasing such a system would incur costs.

Technical strategies to deter cheating include randomising questions, shuffling question and options order, limiting the number of attempts, limiting the time for tests and deferring automated feedback until after the test closes. For more detailed and additional technical strategies readers are referred to documents such as ‘Best Practices for On-Line Academic Integrity’ (55) and ‘Online eAssessment: AMEE Guide No. 39’ (37).

CLOSING

The COVID-19 pandemic may have come with a silver lining (58). Although ERT implies a temporary state of things, it is likely that after the critical period things will not return completely as before (59). OT&L and online assessment are likely to occupy a higher percentage of the future medical curriculum, which can be seen as a positive development for online learning (5, 59). Given this likelihood, medical schools would do well to prepare their physical and human resources and design their OT&L and online assessments using good design principles.

REFERENCES


31. Exam cancellations (Coronavirus) [Internet]. Royal College of Psychiatrists. [cited 29 Apr 2020]. Available from: https://www.rcpsych.ac.uk/training/exams/exam-cancellations


46. Types of online assessments [Internet]. Online assessment. 2020 [cited 3 May 2020]. Available from: https://sites.psu.edu/onlineassessment/gather-evidence/


53. Krueger K. How to catch students cheating on online tests [Internet]. [cited 26 Apr 2020]. Available from: http://mediashift.org/2015/08/how-to-catch-students-cheating-on-online-tests/


