



Preserving Cornerstones of Student's Assessment in Medical Education During COVID-19

Pedro Tadao Hamamoto Filho^{1†}, Angélica Maria Bicudo^{2†} and Dario Cecilio-Fernandes^{3†}

¹ Department of Neurology, Psychology and Psychiatry, Universidade Estadual Paulista (UNESP), Botucatu, Brazil,

² Department of Pediatrics, University of Campinas (UNICAMP), Campinas, Brazil, ³ Department of Medical Psychology and Psychiatry, University of Campinas (UNICAMP), Campinas, Brazil

Keywords: assessment and education, COVID-19, remote (distant) education, medical education, evaluation

OPEN ACCESS

Edited by:

Rhoda Scherman,
Auckland University of Technology,
New Zealand

Reviewed by:

Elena Aurel Railean,
Other, Chisinau, Moldova
Rubia Cobo Rendon,
University of Concepcion, Chile

*Correspondence:

Pedro Tadao Hamamoto Filho
pedro.hamamoto@unesp.br

†ORCID:

Pedro Tadao Hamamoto Filho
orcid.org/0000-0001-6436-9307
Angélica Maria Bicudo
orcid.org/0000-0003-3043-5147
Dario Cecilio-Fernandes
orcid.org/0000-0002-8746-1680

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology

Received: 03 August 2020

Accepted: 10 March 2021

Published: 09 April 2021

Citation:

Hamamoto Filho PT, Bicudo AM and
Cecilio-Fernandes D (2021)
Preserving Cornerstones of Student's
Assessment in Medical Education
During COVID-19.
Front. Psychol. 12:591152.
doi: 10.3389/fpsyg.2021.591152

Since the outbreak of the pandemic caused by COVID-19, there was a significant disruption in our day-to-day life. To control the spread of COVID-19, several measures were taken, including social distancing and isolation. All educational systems were seriously compromised, and medical education, in particular, was heavily hit by those measurements, leading to the challenge of redesigning the current form of teaching. Worldwide, there was a trend of moving all the learning and teaching activities to online. The transition from face-to-face to online activities was not an easy task, especially for low- and middle-income countries (Cecilio-Fernandes et al., 2020). Overall, the transition focuses on the knowledge part of the curriculum and a few medical skills, which were predominantly skills that required non-motor skills, such as communication and clinical reasoning. Although an essential part of the medical curriculum is work-based learning, there is still a need to assess students, as we come to the end of the semester. In this perspective paper, we highlight different methods of assessment and how we make a sound decision about students' progression.

As online teaching has focused on knowledge, it is not fair and feasible to assess students' skills that are taught at bedside, which includes most of the medical skills, attitudes, and competence. Since a proper alignment between learning objectives, strategies, and assessment are vital for any medical curriculum (Prideaux, 2003, 2007), students' assessment should be re-aligned with the new content of teaching. Since the online learning objectives are low in Miller's pyramid, it is only possible to assess students at the level of "knowing" and "knowing how" (Miller, 1990; Wass et al., 2001). Therefore, we should only use assessment tools that are aligned with those two levels.

Different assessment methods can be used to assess the "knowing" and "knowing how" level at Miller's pyramid (Miller, 1990). Although this article will not discuss the various methods, here is a shortlist that may be used in online assessment: multiple-choice questions, essays, short answer, very short answer, open-ended questions, open-book questions, case study, oral examination, and many others. Indeed, the problem is not in the lack of options but in how we can make a fair assessment while assuring that students will not cheat in an online environment.

To answer that question, first we need to understand that end of semester tests, or having two tests a semester, was already lacking. Students often cram before end of semester tests, leading to a low retention of the material. Also, students are often overwhelmed with the number of tests at the end of semester, which forces them to study for some tests while disregarding others. With the transition to online teaching and now assessment many medical educators realized that they should use other types of assessment methods, creating an opportunity to use best practices in assessment. Many curricula focus the decision of pass/fail students on a high-stakes test. High-stakes test refers to the stake of the exam, meaning that if only one test will decide whether you pass or fail, we call it a high-stakes test (Norcini and McKinley, 2007). This type of test leads to students studying for one test, which hampers learning (Wrigley, 2012). Also, this type of test is relatively applicable in

face-to-face education, since all students have to take the assessment at the same time and at the same place, which, in turn, makes cheating unlikely. In an online environment, however, we cannot control the setting, opening the door for cheating. Therefore, many medical educators are looking for different strategies of assessment for an online environment.

If the idea is to keep the high-stakes test, there are two options. First, there are online proctoring and psychometric forensic programs that can help. Those programs make it possible to watch all the students via webcam and more advanced ones also track the gaze of your eyes, making sure that students are only looking at the screen (Weiner and Hertz, 2017). Psychometric forensic refers to track all questions to verify whether there is a change on the item psychometric pattern, during and after the test (Cizek and Wollack, 2017). Those software are expensive and require specialized workers that currently are scarce in Brazil. Second, a computerized adaptive test could be used. The computerized adaptive test refers to an algorithm that will select the next questions based on students' previous performance (Cecilio-Fernandes, 2019; Collares and Cecilio-Fernandes, 2019). This makes the test unique to every student, therefore decreasing the likelihood of cheating. Also, computerized adaptive tests identify students' knowledge gaps more precisely than the traditional format. Finally, students' motivation increases, since they will answer a test that is aligned with their level of knowledge. Although there is free software for computerized adaptive tests, there is a need to have a psychometrician with experience running this type of test. These two options may be feasible for some medical schools, but the majority may not have enough expertise or funding, due to the complexity of both solutions.

From a psychometric point of view, high-stakes tests should provide evidence of validity, reliability, and standardization. Validity refers to whether the test is measuring the proposed construct. Reliability refers to the amount of error in a test and whether the scores are reproducible over time. Standardization refers to how the scores are calculated based on a large sample (American Educational Research Association, 2014). Unfortunately, high-stakes tests in medical schools do not achieve the psychometric rigor necessary to make a fair decision, leading to an unfair decision in many times.

Shifting the paradigm from high-stakes tests to several low-stakes tests seems more realistic to our reality. Actually, rather than being concerned with the pass/fail score, we now move to the original essence of assessment: did students learn? This new paradigm also brings us close to current best guides of assessment, moving toward a system of assessment (Norcini et al., 2018) or programmatic assessment (Prideaux, 2003; Schuwirth and van der Vleuten, 2011; van der Vleuten et al., 2012, p. 14). Both of these refer to using multiple methods of assessment during different points in time. Then, all these different methods and assessment are considered for the decision of the students. Considering multiple assessments has the benefit of changing students' behavior. For example, instead of studying for one test, now they have to study across the semester or year,

which, in turn, is beneficial for students' learning (Cecilio-Fernandes et al., 2017). Also, incorporating multiple methods of assessment gives a clearer evaluation of students as a whole and also decreases the likelihood of cheating, since students will have many tests. It is also possible to include clinical cases discussion, allowing a discussion between students. Using this system opens the door for many possibilities while assuring the psychometric rigor of the assessment program (Bok et al., 2018).

Making a decision on multiple tests also allows to include the formative character of assessment. By formative character, we mean feedback, since studies have shown that students feel demotivated with formative assessment compared to summative assessment. Also, there is a call to reconcile summative and formative assessment, since they are connected at their inception (Lau, 2016). Feedback is one of the most powerful tools for learning (for a review, please see Hattie and Timperley, 2007). Feedback avoids learning an incorrect knowledge by guiding students to the correct knowledge. Feedback also helps learners to identify the relevant knowledge and it motivates students by rewarding correct actions or correcting incorrect actions. Changing from high-stakes tests to low-stakes tests allows medical educators to incorporate feedback in the assessment, increasing students' learning.

Most high-stakes tests are based on closed- or open-ended questions. Using a multiple method of assessments allows other assessment methods to be included. For example, medical educators can now take the advantage of quiz at the beginning of each class, which would help teachers to identify students' knowledge gaps of previous content and also take advantage of the testing effect. Testing effect refers to students who are tested to perform better in a retention test than those who re-studied the material (Roediger and Karpicke, 2006). Also, teachers could include case discussion and group work as part of the decision making. Including different methods of assessment allows for assessing students with a different lens and given feedback on different aspects of students' knowledge.

Noteworthy, online teaching and assessment have mainly focused on knowledge. We have demonstrated that it is possible to make a fair decision on the students' future. However, this may be limited to preclinical training. For the clinical years, it is possible to teach and assess to an extent. Of course, we are going through difficult times, but a discussion about clinical training is necessary. For example, we do not believe that it is possible to teach and assess clinical skills using online tools. Even clinical reasoning that is based on knowledge, the online platform, discussion, and cases are oversimplified compared to real-life scenarios. This discussion is even more important for the 6th year students who are about to graduate. Unfortunately, there is no right answer or a simple solution. Nevertheless, we need to consider whether we should approve students even cutting at least 30% of their practical training, which is perhaps the crucial part of the medical curriculum.

Shifting from face-to-face to online activities may have brought us a unique opportunity to change our knowledge assessment practice. Changing from high-stakes assessment to

multiple low-stakes assessments is an advance to the assessment field, and we hope that this change will last after the pandemic.

AUTHOR CONTRIBUTIONS

DC-F, AB, and PH contributed to the conceptualization, writing original draft, and writing—review and editing. All authors contributed to the article and approved the submitted version.

REFERENCES

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington, DC.
- Bok, H. G., de Jong, L. H., O'Neill, T., Maxey, C., and Hecker, K. G. (2018). Validity evidence for programmatic assessment in competency-based education. *Perspect. Med. Educ.* 7, 362–372. doi: 10.1007/s40037-018-0481-2
- Cecilio-Fernandes, D. (2019). Implementing computerised adaptive test. *Sci. Med.* 29:34432. doi: 10.15448/1980-6108.2019.3.34432
- Cecilio-Fernandes, D., Cohen-Schotanus, J., and Tio, R. A. (2017). Assessment programs to enhance learning. *Phys. Therapy Rev.* 23, 17–20. doi: 10.1080/10833196.2017.1341143
- Cecilio-Fernandes, D., Parisi, M. C. R., Santos, T. M., and Sandars, J. (2020). The COVID-19 pandemic and the challenge of using technology for medical education in low and middle income countries. *MedEdPublish* 9:74. doi: 10.15694/mep.2020.000074.1
- Cizek, G. J., and Wollack, J. A. (2017). *Handbook of Quantitative Methods for Detecting Cheating on Tests*. Washington, DC: Routledge Taylor and Francis. doi: 10.4324/9781315743097
- Collares, C. F., and Cecilio-Fernandes, D. (2019). When I say... computerised adaptive testing. *Med. Educ.* 53, 115–116. doi: 10.1111/medu.13648
- Hattie, J., and Timperley, H. (2007). The power of feedback. *Rev. Educat. Res.* 77, 81–112. doi: 10.3102/003465430298487
- Lau, A. M. S. (2016). 'Formative good, summative bad?'—A review of the dichotomy in assessment literature. *J. Further Higher Educat.* 40, 509–525. doi: 10.1080/0309877X.2014.984600
- Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Acad. Med.* 65, S63–S67. doi: 10.1097/00001888-199009000-00045
- Norcini, J., Anderson, M. B., Bollela, V., Burch, V., Costa, M. J., Duvivier, R., et al. (2018). 2018 Consensus framework for good assessment. *Med. Teach.* 40, 1102–1109. doi: 10.1080/0142159X.2018.1500016
- Norcini, J. J., and McKinley, D. W. (2007). Assessment methods in medical education. *Teach. Teach. Educ.* 23, 239–250. doi: 10.1016/j.tate.2006.12.021

FUNDING

DC-F was partially funded by FAPESP—São Paulo Research Foundation [Young Investigator Grant No. 2018/15642-1]. The views and opinions are entirely those of the authors and the funder had no role in selection of content, decision to publish, or preparation of the manuscript.

- Prideaux, D. (2003). ABC of learning and teaching in medicine. Curriculum design. *BMJ* 326, 268–270. doi: 10.1136/bmj.326.7383.268
- Prideaux, D. (2007). Curriculum development in medical education: from acronyms to dynamism. *Teach. Teach. Educ.* 23, 294–302. doi: 10.1016/j.tate.2006.12.017
- Roediger, H. L., and Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves long-term retention. *Psychol. Sci.* 17, 249–255. doi: 10.1111/j.1467-9280.2006.01693.x
- Schuwirth, L. W. T., and van der Vleuten, C. P. M. (2011). Programmatic assessment: From assessment of learning to assessment for learning. *Med. Teach.* 33, 478–485. doi: 10.3109/0142159X.2011.565828
- van der Vleuten, C. P., Schuwirth, L. W., Driessen, E. W., Dijkstra, J., Tigelaar, D., Baartman, L. K., et al. (2012). A model for programmatic assessment fit for purpose. *Med. Teach.* 34, 205–214. doi: 10.3109/0142159X.2012.652239
- Wass, V., Van der Vleuten, C., Shatzer, J., and Jones, R. (2001). Assessment of clinical competence. *Lancet* 357, 945–949. doi: 10.1016/S0140-6736(00)04221-5
- Weiner, J. A., and Hurtz, G. M. (2017). A comparative study of online remote proctored versus onsite proctored high-stakes exams. *J. Appl. Test. Tech.* 18, 13–20. Available online at: <http://www.jattjournal.com/index.php/atp/article/view/113061>
- Wrigley, W. (2012). A systemic framework for the progress test: strengths, constraints and issues: AMEE guide no. 71. *Med. Teach.* 34, 683–697. doi: 10.3109/0142159X.2012.704437

Conflict 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.

Copyright © 2021 Hamamoto Filho, Bicudo and Cecilio-Fernandes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.