#### Multiple-Choice Assessment in Higher Education: Are We Moving Backward? A Response

Jon Scoresby Ph.D.

Contact info: jscores@gmail.com

Mary Tkatchov points out in her piece, "Multiple-Choice Assessment in Higher Education: Are We Moving Backward," that multiple-choice assessments have their place in the education process and are appropriate for pre-assessment and formative assessment to give students automated and immediate feedback. She also states that multiple-choice assessments are only appropriate for summative assessment "when a broad representation of content knowledge must be measured before more authentic demonstrations of knowledge," but that the goal of assessment in higher education is to gauge students’ “ability to transfer their learning to real-life situations.” I agree that multiple-choice assessments do not provide students with the opportunity to “do” what may be expected of them in their jobs or in the real world; however, authentic demonstrations of skills through assessment is more desirable in some contexts than others.  Rather than adopting a one-size-fits-all assessment philosophy, higher education institutions will need to balance assessment practices with other institutional, faculty, and student needs and priorities.  The type of course and its place in the program sequence, for example, are variables that affect the need for performance and constructed-response assessment and, therefore, the appropriateness of multiple-choice assessments and the predominant assessment method.

A lot of work and research has gone into the knowledge and skill base for item writing within multiple-choice assessment.  For example, over the years researchers have developed and validated multiple-choice item writing taxonomies (Haladyna & Downing, 1989, 2002). Other researchers found that the way an item is written has psychological effects on students, (Wolf, et.al., 1995), and other researchers provide psychometric tools for testing the validity and reliability of items (DeVon, et.al., 2007). Research has clearly shown that student understanding can be assessed at a high cognitive level (Briggs, et al. 2006). When done well and to a standard of tangible quality, multiple-choice assessment can be made career-relevant.  So, if well-crafted multiple-choice items can measure higher order thinking then these assessments could be reasonably used as summative assessment in classes with large populations that are early enough in the sequence where students would not need to worry about having evidence of skills, or being at the level where they would have to perform “on the job.”  As long as the items challenge critical thinking and can provide career-relevant scenarios, predominantly multiple-choice assessments would be appropriate at this level.

As previously stated, the needs of the institution, faculty members, and students need to be considered if multiple-choice assessments, as stated by Tkatchov, are a way to justify larger class sizes for fewer faculty. There is a cost for developing either multiple-choice or performance-based assessments. I remember taking courses in auditoriums with 300 plus students. The cost (time and money) associated with grading open-ended assessments for a class of this size would not benefit the faculty nor the institution.  Specifically, in large scale testing (ex. standardized tests), the rating of open-ended questions would increase costs and time while introducing bias by the human graders (Briggs et.al. 2006). I think the issue is working towards balancing these three factors and finding the priority of this balance should be left to the institution and faculty.

In this day and age of data collection, data analytics are being used to help improve the operational efficiency of colleges and universities, the student learning experience, and student retention rates (van Barneveld et.al., 2012). The automatic and immediate method of data collection from multiple-choice assessments may be another factor in identifying the balance among the needs of the institution, faculty, and student. Although data can also be collected from performance assessments, the automaticity/immediacy of the multiple-choice options is something to consider when looking at all avenues of efficiency. For example, institutions may be more efficient in their course development process by using test scores to immediately identify content areas where students are struggling by reviewing the number of incorrect answers related to a specific topic. This information can then be used to improve the course structure and or instruction from the faculty.

When trying to find efficiencies or the balance among the needs of the institution, faculty, and student, some institutions are breaking up the roles of faculty. Specifically, some institutions are hiring full-time faculty who only teach and other faculty who only evaluate assessments (Newbold, et.al 2017). With this kind of faculty model, performance-based assessments may not be as cost/time prohibitive because one faculty member does not have to both teach and grade. In these cases, the faculty who are able to focus on teaching, focus on their needs in ways that not only improves their ability to help students learn, but also help the institution create and provide a good service to students.  Those who only evaluate assessments are able to focus on the feedback given to students in an effort to help them improve.

Tkatchov also talks about the value of the performance assessments in that students have the opportunity to not only learn and improve writing skills, they come away with an artifact that shows their level of learning and understanding beyond that of a test score. This point is powerful because multiple-choice assessments are unable to provide students with an artifact that shows transferability of knowledge and skills. Some may be concerned that multiple-choice assessments may not be able to show or measure students’ understanding as well as open-ended performance assessment can. According to Briggs et.al. (2006), well-written multiple-choice items can be linked to students’ cognitive development and measure student understanding. So going back to the idea of finding the right balance, the automaticity and immediacy of feedback to students and faculty from well-designed items can be a powerful tool for improving the learning experience.  Well-written assessment items may play a role in finding the right balance to meet the needs of the institution, faculty, and students.

I previously stated that the type of course and its place in the program sequence may affect the type of assessment implemented. By the end of students’ programs, it would seem more appropriate to provide students with opportunities to demonstrate their knowledge and skills in real-world context and be able to graduate with evidence of their knowledge and skills, thus performance-based assessments may be better suited as the predominant method of assessment.  I do not know if a perfect balance will be ever be found with multiple-choice and performance-based assessments and the multiple needs within an institution of higher education, but because multiple-choice assessments will not be going away anytime soon, maybe the answer lies within striving to improve faculty skills in item writing.

References

Briggs, D., Alonzo, A., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment,* *11*(1), 33-63. doi:10.1207/s15326977ea1101\_2

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J. et al. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship,* *39*(2), 155-164. doi:10.1111/j.1547-5069.2007.00161.x

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education,* *15*(3), 309-333. doi:10.1207/s15324818ame1503\_5

Haladyna, T. M., & Downing, S. M. (1989). Validity of a taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education,* *2*(1), 51-78. doi:10.1207/s15324818ame0201\_4

Newbold, C., Seifert, C., Doherty, B., Scheffler, A., & Ray, A. (2017). Ensuring faculty success in online competency-based education programs. *The Journal of Competency-Based Education,* *2*(3). doi:10.1002/cbe2.1052

van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). Analytics in higher education: Establishing a common language. *Educause Learning Initiative*, 1, 1–11.

Wolf, L. F., Smith, J. K., & Birnbaum, M. E. (1995). Consequence of performance, test, motivation, and mentally taxing items. *Applied Measurement in Education,* *8*(4), 341-351. doi:10.1207/s15324818ame0804\_4