

Prohibiting Students from Asking Questions during Exams: A Guideline for Promoting Fairness and Preserving Score Validity

Kenneth D. Royal ■ Karen R. Munana ■ Jennifer A. Neel

ABSTRACT

Many medical and professional programs implement policies that prohibit students from asking questions during examinations. The reasoning behind these policies remains unclear to some, as there is a lack of literature addressing this topic. The purpose of this article is to present the rationale behind such policies and to discuss why these policies may help promote fairness and preserve score validity.

Key words: academic affairs, assessment, curriculum, disabilities, educational methods, SKAs, student affairs

At the North Carolina State University (NCSU) College of Veterinary Medicine (CVM), we recently updated our guidelines for administering examinations within the CVM. The guidelines were created based on “best practices” for professional testing in the fields of examination security, administration, and psychometrics. The guidelines were formally reviewed and approved by the Faculty Committee on Academic Performance and Student Conduct (FCAPSC), which consists of several faculty members, administrators, and students in the doctor of veterinary medicine (DVM) program. The revised guidelines were presented at a college-wide faculty meeting in the fall semester of 2015, and went into effect in the spring semester of 2016. While faculty were generally supportive of most updates and changes, one particular item was troubling for some faculty members. Specifically, our new guidelines state the following:

No questions will be answered during the exam (e.g., clarification, rewording, interpretation, suspected typos, etc.). Any questions with typos or other errors will be assessed post exam for potential elimination. An in-test mechanism for voicing challenges/concerns regarding exam questions will be provided and will be clearly identified at the start of the exam.

The guidelines go on to list a variety of mechanisms for voicing a challenge or concern, depending on whether the examination is conducted electronically or via Scantron.

After the guidelines went into effect and faculty were asked to adopt standardized examination practices and procedures, several challenges were presented to the “no questions” guideline. Faculty who insisted on permitting

students to ask questions during examinations cited the following reasons: (1) Allowing questions may help the instructor identify which items are confusing, so efforts can be made to improve the item for future exams; (2) Allowing questions ensures that students know exactly what the instructor is asking and allows for testing of the material rather than students’ interpretations of the instructor’s question; (3) Not allowing questions results in increased grading time, as the instructor would need to respond to written challenges made by students while taking the exam; and (4) Questions asked during an exam may provide an additional opportunity for the faculty to learn what material is not well understood by the students and why. In our view, each of the aforementioned arguments could be easily addressed in other ways. For example, concerns about item quality can largely be mitigated if one simply follows “best practices” of item writing, which are routinely discussed and made available to our faculty. Second, allowing students a mechanism to flag potentially problematic items will still provide faculty feedback about any items that are potentially confusing. Third, grading time would be largely unaffected as all item statistics and student-rendered flags are evaluated after each examination as part of standard assessment practices in the college. Finally, as it pertains to measuring learning, faculty can simply review students’ performance at the item/content level and identify sufficient clues about which material was not learned well and requires additional clarification. Of course, faculty could simply ask students at the beginning of the next class session if any material remains unclear. In our view, none of the arguments for permitting students to ask questions outweighed the disadvantages of this practice. Thus, the

purpose of this article is to discuss guidelines and policies in which students are not permitted to ask questions during examinations and to present the rationale behind these necessary, albeit sometimes controversial, practices.

THE ROLE OF STANDARDIZATION IN PROMOTING FAIRNESS AND SCORE VALIDITY

Most examinations conducted in medical and health professional programs are moderate-to-high stakes for students.¹ The *Standards for Educational and Psychological Testing*² essentially states that the greater the stakes for examinees, the greater the amount of evidence is necessary to support the valid inferences one might make about a set of scores. To carry out assessments that are both fair for examinees and capable of yielding valid scores, assessment professionals rely heavily on standardized practices. As noted by Royal and Hecker,³ test scores are typically affected by three primary factors: (1) the instrument, (2) the examinees, and (3) the conditions under which the examination was administered. When instruments and the conditions for which examinations are administered are standardized, the error stemming from these sources is mitigated. Any variation in examinees' scores can be more directly attributed (with greater certainty) to actual differences in student ability. Thus, the importance of standardized processes cannot be understated. It was with this backdrop in mind that we identified five reasons to prohibit students from asking questions during examinations. The following section presents the five reasons and discusses how each affects fairness and validity concerns.

RATIONALE FOR PROHIBITING STUDENTS FROM ASKING QUESTIONS DURING EXAMS

In recent years, the number of students with disabilities has increased exponentially.⁴ The American Disabilities Act (ADA) mandates that persons with documented disabilities be provided reasonable accommodations during testing scenarios.⁵ Such accommodations often include extended time on tests, providing quiet/private testing rooms (i.e., distraction and noise-reduced environments), and so on. Virtually all medical and health programs now have students with documented disabilities, thus it is necessary to consider how efforts to accommodate students with disabilities might affect other students as well. Because students with accommodations are often tested in a different physical location (to reduce noise and other distractions), students who ask questions in the primary testing setting will be privy to answers that students in a different location may not hear. This lack of standardization disadvantages any student who tests elsewhere (e.g., students with accommodations testing in a private room or students testing at a later time due to an excused absence).

Second, students complete exams at different speeds and typically are free to leave the testing environment immediately upon completion. This raises the question of what happens when someone asks a question after

some students have already completed the exam and left the testing environment. In such instances, students who complete an exam early would be disadvantaged, and those still present in the test setting area would be advantaged. Certainly, it would be possible to require students to remain in the testing area until everyone has completed the exam, but such a policy would likely be unpopular with students and would invariably result in increased threats to exam security. Thus, when weighing the advantages and disadvantages of requiring students to remain in the testing area versus simply prohibiting students from asking questions, it is clear that prohibiting questions would be more practical and effective.

Third, asking questions during an examination may disturb some students. Professional testing providers such as Prometric and Pearson VUE go to great pains to create a distraction-free environment, as distractions such as excessive noise can affect examinees' and can ultimately result in invalid scores. While asking a single question before the entire group of students during an examination might not seem particularly egregious, instances in which students ask multiple questions could become particularly distracting and affect some students' concentration, possibly leading to increased anxiety. Thus, it is possible that students who have trouble concentrating or who experience anxiety (without accommodations) might be unduly disadvantaged by any talking during an exam. Of course, sometimes students approach faculty in private and ask questions. When this occurs, it often creates a dilemma for the instructor of whether or not to make a general announcement to the entire class. If the instructor answers only the single student's question, then all other students may potentially be disadvantaged. However, if the instructor decides to answer the question and announce the answer to the entire class, it does mitigate the fairness issue (at least for those present in the room), but it may disturb some students who did not need clarification.

Fourth, there is a concern that faculty may simply "give too much away" when responding to a question. By their very nature, faculty members are typically warm, approachable, and helpful individuals. When students ask for help, most educators will find it difficult to ignore their request. Thus, when students are permitted to ask questions during an exam, it essentially puts instructors in a position in which they might feel somewhat obligated to respond. In their efforts to be helpful, some instructors might inadvertently provide information (e.g., a subtle hint, a cue intended to help students recall something) that leads examinees to the correct answer. When this occurs, students' responses to the item are now contaminated with error that may ultimately affect both the accuracy and meaning (validity) of their scores.

Finally, there is concern that when instructors present an item differently (e.g., rephrase the question, provide additional information) it may alter the meaning of the item and ultimately how the item functions statistically. Previous research has shown that very trivial changes to an item can affect how examinees respond, thereby affecting the item's psychometric functioning.⁶ Thus, it is reasonable to assume that any alteration of an item, including

a verbal response, may affect how students respond and ultimately affect the integrity of the item. Further, when instructors evaluate various psychometric indicators of item quality (e.g., measures of difficulty, discrimination, etc.), they might forget about providing alternative or additional information to students for a particular item or items. This could result in several unintended consequences, such as identifying an item that “jumped” considerably in terms of difficulty (which typically is indicative of previous item exposure), or identifying an item that is now so “easy” that it is rendered ineffective. In either scenario, instructors may be tempted to remove the presumably problematic item from future examinations, thus costing them valuable time and energy as they generate a new (and robust) item on similar content. In any instance, presenting an item in any other way than what is presented on the assessment could introduce unwanted error and affect score accuracy and meaning.

ADDITIONAL CONSIDERATIONS

We have presented five reasons for adopting a guideline and/or policy that prohibits students from asking questions during an examination. While the guideline is intended to promote fairness and ensure score validity, it is important to recognize that faculty, on occasion, do present students with items that are ambiguous or of questionable quality. That said, it is critical that institutions adopting a similar “no questions during exams” guideline/policy provide students some mechanism for flagging items that may be problematic. For examinations administered electronically, most assessment programs have a flagging feature for examinees to call attention to a potentially problematic item. Some software will also allow examinees to comment on the nature of the problem. In the case of paper-and-pencil and Scantron examinations, students should also be provided an opportunity to express their concerns about a potentially problematic item. We suggest creating a designated section for comments on either the examination or Scantron form. We recommend faculty promptly review any items that were flagged to ensure item quality and score integrity. This practice also has the advantage of capturing student concerns after the exam, as opposed to having to make note of questions that students asked during the exam or trying to remember what problems occurred.

Finally, there is an old adage that says “prevention is the best medicine.” This general truth certainly applies here insofar as instructors should be proactive and take the necessary steps to ensure that their examination items are clear and to assess what they are intended to measure. Abundant resources are available to help faculty construct sound examination items. We recommend that faculty see Haladyna and Rodriguez⁷ for an excellent primer.

REFERENCES

- 1 Royal KD, Hedgpeth MW. A novel method for evaluating examination item quality. *Int J Psychol Stud*. 2015;7(1):17–22. <http://dx.doi.org/10.5539/ijps.v7n1p17>.
- 2 American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association; 2014.
- 3 Royal KD, Hecker KG. Understanding reliability: A review for veterinary educators. *J Vet Med Educ*. 2016;43(1):1–4. <http://dx.doi.org/10.3138/jvme.0315-030R>. Medline:26560547
- 4 Nelson JM, Whipple B, Lindstrom W, et al. How is ADHD assessed and documented? Examination of psychological reports submitted to determine eligibility for postsecondary disability. *J Atten Disord*. 2014. <http://dx.doi.org/10.1177/1087054714561860>. Medline:25534434
- 5 Lovett BJ. Testing accommodations under the Amended Americans with Disabilities Act: the voice of empirical research. *J Disabil Policy Stud*. 2014;25(2):81–90. <http://dx.doi.org/10.1177/1044207312469830>.
- 6 Strickland B, Fisher M, Knobe J. Moral structure falls out of general event structure. *Psychol Inq*. 2012;23(2):198–205. <http://dx.doi.org/10.1080/1047840X.2012.668272>.
- 7 Haladyna TM, Rodriguez MC. *Developing and validating test items*. New York: Routledge; 2013. <http://dx.doi.org/10.4324/9780203850381>.

AUTHOR INFORMATION

Kenneth D. Royal, PhD, MEd, is Assistant Professor of Educational Assessment & Outcomes, North Carolina State University, College of Veterinary Medicine, Department of Clinical Sciences, 1060 William Moore Drive, Raleigh, NC 27607 USA. Email: kdroyal2@ncsu.edu.

Karen R. Munana, DVM, MS, is Professor of Neurology, North Carolina State University, College of Veterinary Medicine, Department of Clinical Sciences, 1060 William Moore Drive, Raleigh, NC 27607 USA. Email: karen_munana@ncsu.edu.

Jennifer A. Neel, DVM, is Assistant Dean for Education and Associate Professor, North Carolina State University, College of Veterinary Medicine, Department of Population Health and Pathobiology, 1060 William Moore Drive, Raleigh, NC 27607 USA. Email: jennifer_neel@ncsu.edu.