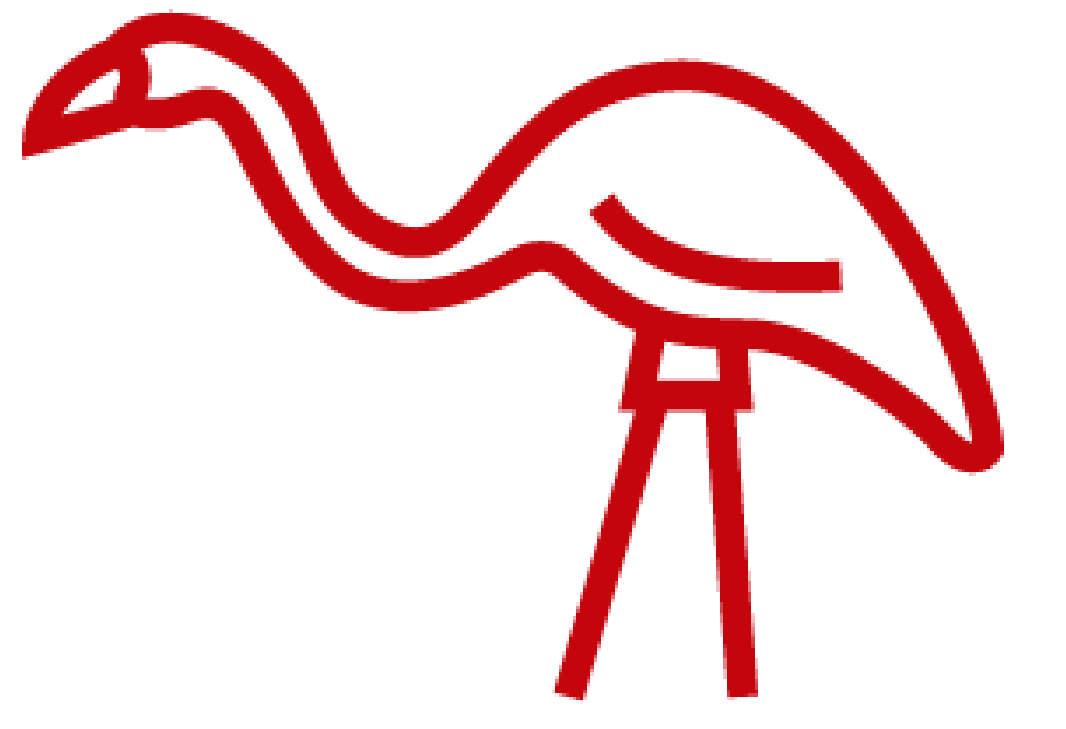


Operationalizing “High Quality” in Short Answer Assessment Questions Across a Pharmacy School Division



Kate Rotzenberg, PharmD, MBA, BCPS and Lucille M. Abrams, MS
University of Wisconsin-Madison School of Pharmacy

Background

Best practices have been published for the construction of short answer assessment questions but there is no tool available to retrospectively measure the quality of these questions in a systematic manner.¹⁻⁴

Methods

The authors conducted a literature review on short answer question construction best practices and used these to create an initial quality evaluation tool. Two rounds of testing, discussion, and revisions occurred before arriving at the final tool (20 items). The final tool was then applied to all short answer exam questions in required Doctor of Pharmacy courses within the Social & Administrative Sciences division during the 2022-23 academic year as part of a quality assurance initiative. This project was determined to not be human subjects research.

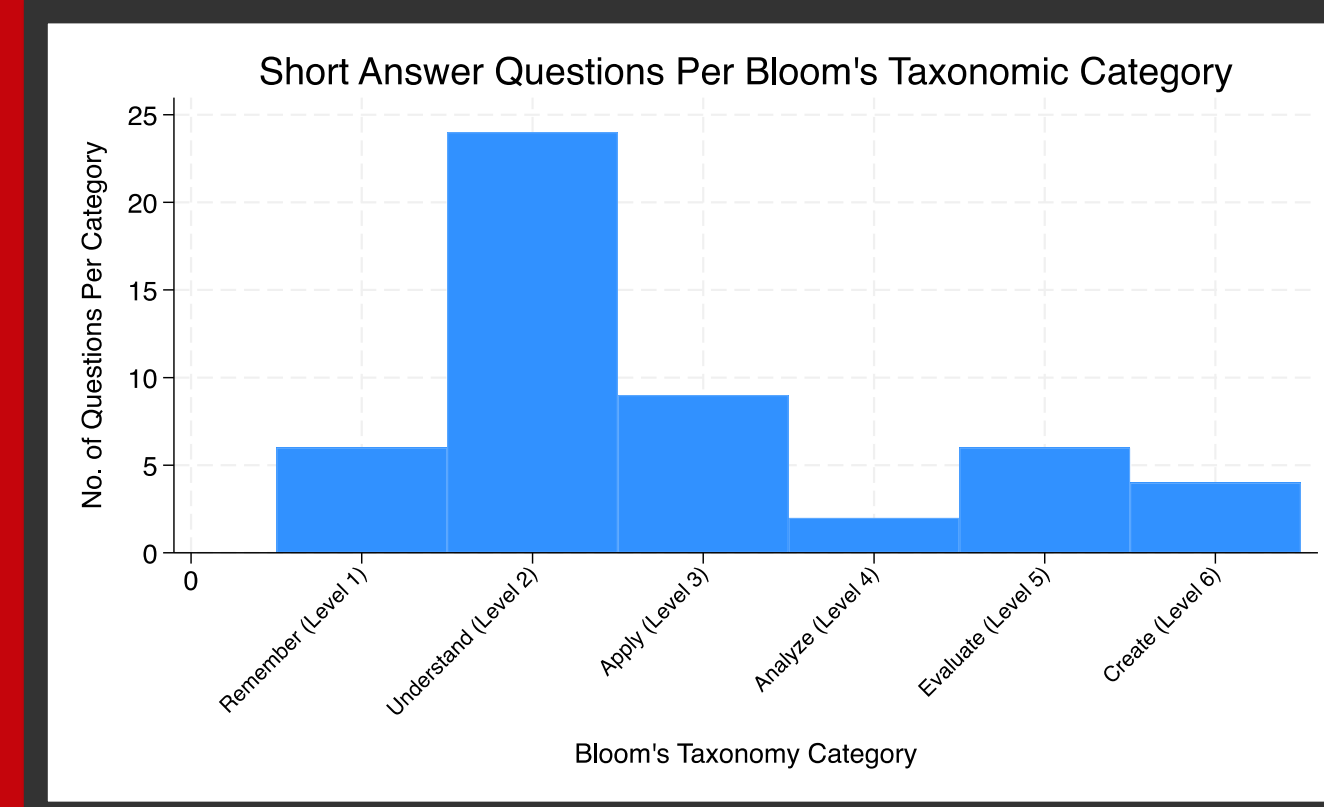
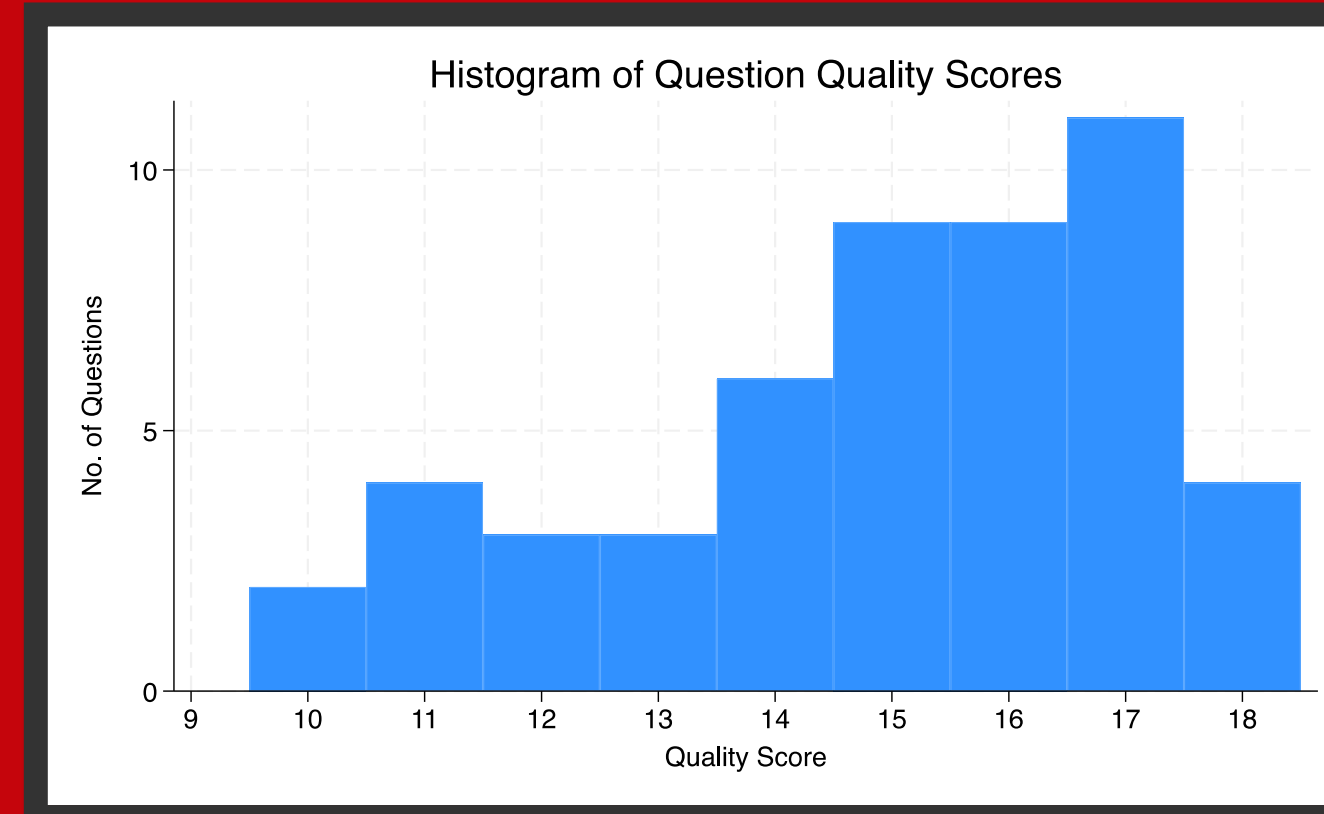
Results

Inter-Rater Agreement

Inter-rater agreement using the final tool on a sample of 26 assessment questions was 88.89% prior to discussion and achieving consensus. The areas most frequently contributing to disagreement were whether key words were properly emphasized with bolding or underlining in a question, which details in a question were considered extraneous, and whether the expected length for written answers was stated clearly.

Adherence to Best Practices

- 51 questions assessed
- Sample mean 14.92 ± 2.23 (maximum score 20, range 12-18)
- See table for the five **most** and five *least* frequently used best practices across the division
- Most questions were written at the “Understand” level of Bloom’s taxonomy



Best Practices Assessed

Domain	Criteria
Cognitive Level	<ul style="list-style-type: none"> • Bloom's taxonomy level of “applied” or higher (42%) • Prompt and learning objective in same level of Bloom's taxonomy (51%)
Clarity in Wording	<ul style="list-style-type: none"> • Prompt uses directive verb • Scope of task clearly outlined (94%) • Prompt positively worded (96%) • All tasks required (no optional tasks) • Prompt uses appropriate vocabulary • Prompt avoids extraneous detail • <i>Prompt includes expected depth of response (12%)</i> • Total point value provided (98%)
Formatting	<ul style="list-style-type: none"> • Prompt is separate from data, if provided • Tasks are numbered/indented, if more than two (90%) • <i>Key words are differentiated (59%)</i>
Scoring	<ul style="list-style-type: none"> • Model answer provided (84%) • Model answer aligns with learning objective • Model answer clearly worded • Model answer fully addresses question • Rubric provided • Rubric addresses all tasks • <i>Tasks are independent or graded forward (53%)</i>

Challenges in Converting a Prospective Checklist to a Retrospective Evaluation

- Inherently subjective process (e.g., what is clear, appropriate, extraneous?) for retrospective evaluators attempting to view question from the perspective of a student
- Retrospective evaluators are not content experts in all areas covered by the short answer questions
 - Limited in ability to determine alignment with learning objectives and clarity of model answers
 - Utility of non-expert peer review as a general practice may be limited to surface-level issues (e.g., formatting, identifying cognitive level mismatch, ensuring rubric matches tasks)
- Cannot meaningfully correlate the quality of the question with student performance
 - Many intervening events occur between writing the question and final scores (e.g., conversations between instructor and grader to resolve clarity issues)
 - Iterative scoring of short answer questions to ensure consistency is normal and expected⁴ but confounds results.

References

- 1) Nguentan D, Gruenberg K, Shin J. Should multiple-choice questions get the SAQ? Development of a short-answer question writing rubric. *Curr Pharm Teach Learn.* 2022;14:591-596.
- 2) Rudolph MJ, Daugherty KK, Ray ME, Shuford VP, Lebovitz L, DiVall MV. Best practices related to examination item construction and post-hoc review. *Am J Pharm Educ.* 2019;83(7):7204.
- 3) Hauer KE, Boscardin C, Brenner JM, van Schaik SM, Papp KK. Twelve tips for assessing medical knowledge with open-ended questions: designing constructed response examinations in medical education. *Med Teach.* 2020;42(8):880-885.
- 4) Hogan TP, Murphy G. Recommendations for preparing and scoring constructed-response items: what the experts say. *Appl Meas Educ.* 2007;20(4):427-441.

*Best practices in **bold** were the five most frequently met; best practices in *italics* were the five least frequently met.