



High-Stakes Exam Review Policy (Midterms/Finals)

Category:	3-Assessment and Evaluation
Curriculum Phase:	Preclerkship
Author/Authoring Body:	Exam Review Subcommittee
Responsible Unit:	Curricular Affairs
Policy Contact:	Assessment & Evaluation Office
Next Review Year:	2025

Summary

This policy provides guidelines for a block/course director in consultation with the Exam Review Subcommittee (ERS) of the Tucson Educational Policy Committee to conduct an item analysis of multiple-choice items (MCIs) on high-stakes exams. These guidelines ensure a consistent practice for managing MCI outcomes across blocks and courses.

Related LCME Standard: 9.4 Assessment System

A medical school ensures that, throughout its medical education program, there is a centralized system in place that employs a variety of measures (including direct observation) for the assessment of student achievement, including students' acquisition of the knowledge, core clinical skills (e.g., medical history-taking, physical examination), behaviors, and attitudes specified in medical education program objectives, and that ensures that all medical students achieve the same medical education program objectives.

Applicability & Scope

This policy applies to high-stakes exams (i.e., midterm and final exams) in the Preclerkship Phase of the program leading to the MD degree.

Definitions

MCIs – Multiple-choice items

Distractors – Incorrect answer option for a multiple-choice item

High-Stakes Exam – Includes block and course midterm and final exams in the Preclerkship Phase

Item Analysis – Within psychometrics, this is a review process for improving the reliability of an exam by evaluating student responses to multiple-choice items and assess the quality of those items and the test overall

Item Statistics – Measurements considered in an item analysis of individual multiple-choice items to ensure they are at an appropriate difficulty level and discriminate well between overall high- and low-performers on the exam

Policy Statement

I. Guiding Principles

- A. Provide the Exam Review Subcommittee (ERS) with consistent guidelines for evaluating the performance of exam items and formulating mitigation strategies for recommendation to the block/course director.
- B. MCIs on high-stakes exams where $\leq 30\%$ of students selected the single best/correct distractor indicate that the items require closer examination.
- C. Difficult items are allowed on the exam if the question is fair, i.e., the item can be reliably traced to a clear learning objective and sufficient coverage of the content in the instructional session(s).
- D. Do not penalize students who correctly answered a question because of an exam design flaw.
- E. Do not reduce the points available on an exam by dropping an item(s) and effectively increasing the weight of the remaining items. Instead, consider “neutralizing” the question.
- F. Mitigate grade inflation, i.e., in an instance when a student who correctly answers the question also receives bonus points, only a small number of students should fall in this category.

Guidelines

For the three types of item analysis findings enumerated below, the following guidelines will be considered when calculating the final exam scores on MCIs. These are guidelines and the ultimate decision for managing these item outcomes is left to the discretion of the block/course director in consultation with the ERS.

- A. Flagging questions according to item performance statistics. All items that fit the following item analysis criteria will be flagged:
 - 1) Item difficulty is < 0.9 (fewer than 90% of students selected the correct item) AND point biserial is < 0.15 .
 - 2) Item difficulty is < 0.60 .
- B. Handling questions when student performance indicates “drop” consideration:
 - 1) If $\leq 40\%$ of students selected the single best/correct distractor, and the question is reliably traced to a clear learning objective and sufficient coverage of the content in the instructional session(s): **Keep the question as written with no adjustments to the calculation of the final exam score.**
 - 2) If $\leq 40\%$ of students selected the single best/correct distractor, but the question cannot be traced to a clear learning objective or sufficient coverage of the content in the instructional session(s): **Set all distractors as “correct,” and award a bonus point for a correct response.**

- 3) If there is no verifiable correct answer or insufficient coverage of the content was provided: **Set all distractors as correct and maintain denominator in the calculation of the final exam score.**
- C. Making recommendations for other exam item performance outcomes, whether numerically flagged or identified by ERS members and/or the block/course director:
- 1) If two or more distractors are found to be equivalent, accept each equivalent distractor as correct.
 - 2) If wording or grammar choices are identified as misleading, ERS and/or the block/course director may identify outcomes as described under Guideline B or suggest elimination or revision of the item on future exams.
 - 3) If images are found to be of poor quality, indecipherable, or misleading, ERS and/or the block/course director may identify outcomes as described under Guideline B or suggest elimination or revision of the item on future exams.
 - 4) If items are found to have significant technical flaws as outlined in the National Board of Medical Examiners guidelines for constructing written test questions, ERS and/or the block/course director may identify outcomes as described under Guideline B or suggest elimination or revision of item on future exams.

****Additional recommendations and/or appeal of specific exam questions are made in consultation with ERS and the block/course director.***

Revision History

7/26/2023: The Tucson Educational Policy Committee approved the policy.