Assessment System – Continuous Quality Improvement of Assessment Process

Rationale

Continuous quality improvement is needed throughout the assessment processes in the Foundational Phase of the curriculum (Blocks 1-4). This procedure will outline a process to improve assessments and improve the vertical and horizontal alignment of content within assessments and provide central oversight by the Executive Curriculum Committee.

Scope

This procedure applies to multiple-choice assessment items within the Foundational Phase of the Curriculum (during Blocks 1-4). This policy outlines responsibilities of Office of Medical Education staff, administrators, course directors and the Office of Data Analysis and Strategy.

Definitions

Point-biserial correlation: a correlation coefficient that discriminates between high scoring and low scoring students. Correlations coefficients range from -1.0 (perfect negative discrimination) to 1.0 (perfect positive discrimination).

CQI: continuous quality improvement

Exam autopsy: a face-to-face post-exam meeting scheduled to occur after each exam in the Foundational Phase of the curriculum in which reviewable test items will be peer reviewed.

Item difficulty: the percent of overall test takers that selected the correct response for a given item. Item difficulty is measured by a p-value.

Reviewable Items Exam: assessment created in testing software that includes all items that have been identified for review. This exam allows for Course Directors to amend questions for future iterations.

Multiple-choice items: for the purposes of this procedure, MC items must meet the standards outlined by the BSOM Testing Procedure as approved by the ECC, including:

- Utilizing single best response format
- Avoiding use of true/false (C-type, K-type, and X-type) questions
- Avoiding negatively phrased A type questions- for example, all of the following are correct except or which of the following statements is not correct
- Writing options that are grammatically consistent and logically compatible with the stem.
- Avoiding use of imprecise terms (usually, frequently, often, commonly, most of the time, almost never)
Reviewable test items: any multiple-choice question that returns a p-value greater than or equal to 0.90, p-value less than or equal to 0.40, a point-biserial correlation below .20, or any item that is successfully challenged.

p-value: measure of item difficulty where lower values indicate more difficult items while higher values indicate easier items (range: 0-1). For this procedure, easier items are defined as having a p-value greater than or equal to 0.90; more difficult items are defined as having a p-value of less than or equal to 0.40.

Responsibilities

Multiple people contribute to the continuous improvement of test items.

Course Director: Course directors are responsible for submitting high quality test items. Following each exam, the course director should review any items that are identified via item analysis or test challenge process. They should seek to correct/amend/archive questions based on identified weaknesses in item design. Each course director that submits questions for a given exam block will receive an item analysis to review, submit information to OME regarding questions that were successfully challenged, and participate in the “exam autopsy” (face-to-face posttest meeting) to improve items for future iterations.

Testing Administrator – Office of Medical Education: Testing administrator is responsible for the creation of an item analysis that will identify any questions that result in a p-value greater than or equal to 0.90 or a p-value less than or equal to 0.40. Test administrator will communicate with Course Directors regarding items that will be placed on the Reviewable Item Exam. Test administrator will create the mock exam following the secure exam review and test challenge period prior to the exam autopsy. Testing administrator will attend autopsy meetings and gather data to be reported.

Curriculum coordinator – Office of Medical Education: Curriculum coordinator is responsible for scheduling the “exam autopsy” to occur approximately two weeks after each exam in the Foundational Phase of the curriculum (after exams during Blocks 1-4).

Data analyst – Office of Data Analysis and Strategy: Data analyst is responsible for assisting with the analyzing and reporting of CQI data collected from this process. Data analyst will communicate with the Testing Administrator and Assistant Dean of Assessment and Outcomes to create reports/dashboards that are updated following each exam block. Additionally, aggregate data from the individual blocks and Foundational Phase will be reported out annually.

Assistant Dean of Assessment and Outcomes – Office of Medical Education: Assistant Dean is responsible for reviewing the item analysis produced by the testing coordinator after each exam. The Assistant Dean is also responsible for providing oversight of the CQI process related to testing and reporting progress to the ECC and annual curriculum retreats.
Procedures for Implementation and Review

Post-exam period for review of test item quality

Following the administration of a Foundational Phase exam, several processes occur:

- Secure exam review – specific time dedicated for students to review exam questions, reflect on rationales and reasons for correct answers, and provide an educational opportunity to gain insight into tested concepts. Secure exam reviews should be part of the block calendar approved by the ECC.
- Item analysis – Testing Administrator identifies questions that should be reviewed and reports individual items to the appropriate Course Directors.
- Test item challenge process – after the secure exam review students have up to two business days to submit challenges.
- Assessment quality assurance – Course Directors review item analysis report and student challenges to determine if credit for items needs to be amended. All amendments will be reported from the Course Director to the Testing Administrator in order to create a Reviewable Item Exam for use during the “exam autopsy”.
- Mock exam creation – Testing administrator will create a mock exam five business days after the secure review. This mock exam will be used in the face-to-face meeting (“exam autopsy”).
- Exam autopsy - A face-to-face meeting will be scheduled to occur approximately 6-10 days following the secure review. Meetings should be scheduled prior to the start of the block and will be two hours in length. All reviewable test items will be discussed during the exam autopsy. Items that require amendment or retiring will be modified appropriately and stored for future iterations at this meeting.

Reporting of CQI results

OME captures data on the number of test items that fall into each of the following categories:

- p-value greater than or equal to 0.90
- p-value less than or equal to 0.40
- point-biserial less than .20
- items successfully challenged

This data will be collected for each of the participating courses and aggregated for the block and phase. Data Analyst working with OME will develop a log of updated items to track successful improvements to items over time.

Data to be reported to ECC

- categories for reviewable test items (based on p-values or successful challenges)
• categories for items amended at the meeting (questions with multiple correct answers; questions with all answers accepted/question dropped; questions with typographical/grammatical errors; questions in an improper format)
• participation from course directors

ECC will be updated following each iteration of exam autopsies with the data above. Additionally, aggregate data will be presented annually at the Foundational Sciences Course Directors Retreat.

Schedule for Review and Update

This procedure is reviewed and revised by the Executive Curriculum Committee as necessary, but at least every three – five years.

Related Procedures

SOP 9.8.2 – BSOM Testing Procedure
SOP 9.4 – Assessment System - Formative Assessment and Feedback
SOP 9.5 – Assessment System – Narrative Assessment

Applicable Laws, Regulations & Standards

LCME 8.3 Curricular Design, Review, and Revision/Content Monitoring
LCME 9.0 Teaching, Supervision, Assessment, and Student and Patient Safety
LCME 9.4 Assessment System