

The CAS program enables medical school faculty to build high-quality examinations for students tailored to local curricula. This document provides guidance for interpreting feedback from CAS exams. Importantly, NBME neither sets nor recommends a "passing" score, and CAS exams should generally be used in conjunction with other indicators of student performance to determine grades.

Customized Examination Scores

Performance is reported as percent correct scores for all score categories on the exam, including the total test. Percent correct scores represent the percentage of questions answered correctly. A score category must have at least 10 questions to be reported. When a statistic cannot be computed because business rules are not met, a single dash (–) is displayed.

Precision of Scores

The total test score typically provides the most reliable summary of performance, but all assessment results include some level of imprecision. When there are at least 10 examinees in an administration, a reliability coefficient and standard error (SE) are reported for each score category. The reliability coefficient indicates the expected consistency of scores, whereas the SE indicates how much examinees' scores might vary across administrations. Specifically, if an examinee tested repeatedly under the same conditions on a different set of items covering the same content, without learning or forgetting, their total percent correct score would fall within one SE of the original score two-thirds of the time.

In some administrations, the estimated reliabilities may be negative due to instability associated with small sample sizes or questions measuring more than one construct. In these cases, the SE is reported as not available (NA) to avoid misinterpretation.

Score and Performance Feedback

Group Performance Report

The report presents individual scores and several group-level results for the current administration. It helps faculty examine score distributions and understand how the group performed across score categories and individual questions. This report has multiple sections, which are described below.

Percent Correct Score Descriptive Statistics

This section presents group-level percent correct descriptive statistics for each score category. For every category, the number of questions is reported. When there are at least 5 examinees in an administration, the observed score range (minimum and maximum), mean, and standard deviation are also provided. When an administration includes at least 10 examinees, reliability coefficients and SEs are reported as well.

Roster of Percent Correct Scores

This section lists each examinee's percent correct scores for the total test and each additional score category.

A downloadable Excel file ("Score Roster") provides the same information and is available for every administration.

Total Percent Correct Score Frequency Distribution

This section summarizes the distribution of examinees' total percent correct scores within the administration group. A

table reports the frequency and cumulative frequency for each observed score, along with the corresponding percentages. The cumulative percentages represent percentile ranks, indicating the proportion of scores that fall at or below a given value.

Source Comparisons

This section provides additional score interpretation context by summarizing performance relative to the source exam from which questions were selected (e.g., United States Medical Licensing Examination® Step exam or NBME® Subject Exam). For each score category, it presents the total number of questions, the number of questions from the source exam, the administration group's average percent correct (if there were at least 5 examinees in the administration), the corresponding average percent correct from the source exam, and the difference between these averages (also if there were at least 5 examinees in the administration).

To ensure that these comparisons are meaningful, source exam information is included only when at least 51% of the questions in a score category come from that exam and when the category contains at least 10 questions from the same source. These comparisons can provide useful context; however, they should be interpreted cautiously. Differences in averages may reflect variation in examinee populations, instructional environments, or exam purposes rather than true differences in performance.

(Group) Question-Level Feedback

This section provides question-level performance information. For each question, it presents the source exam, the administration group's percent correct (if there were at least 5 examinees in the administration), the corresponding percent correct from the source exam, and the difference between these values (also if there were at least 5 examinees in the administration). When an administration includes 40 or more examinees, this section also provides a question discrimination statistic, indicating how well, on average, an item on the examination discriminates between high-scoring and low-scoring examinees within the administration.

A downloadable Excel file ("Item Analysis") provides the same information and is available for administrations with at least 5 examinees. A separate Excel file ("Question Level Feedback All Examinees") shows how each examinee answered each question, along with important question-level results at the group level, and is available for administrations with at least 5 examinees.

Examinee Performance Report

The report is provided for each examinee in the current administration. It presents an examinee's percent correct score for the total test and each additional score category, along with the number of questions in each category. If at least 51% of the questions in a category originate from the same source exam and that category contains at least 10 such questions, the report also includes a historic percent correct, providing an approximate performance benchmark from when questions in that category were administered on NBME exams. Although this historic value is the same as the "source % correct" presented in the Group Performance Report, the source exam is not identified on the Examinee Performance Report. Faculty may choose whether to disclose the source exam to examinees.

(Examinee) Question-Level Feedback

This section provides question-level information for each examinee, including each question's content topic, content description, and whether it was answered correctly or incorrectly.

Item Analysis (IA) Application

The IA application allows faculty to review detailed question-level performance for the current administration, including full question text and images. For each question, faculty can examine content details, score category membership, and the percentage of examinees who selected each answer option. They can also review the percent correct for the current administration ("school difficulty," available if there were at least 5 examinees) and the question discrimination statistic ("school discrimination," available if there were at least 40 examinees). In addition, the application displays the source exam and its corresponding source-exam's difficulty (percent correct).

The IA application is only accessible by faculty. For CAS Proctored exams, it cannot be shared with examinees under any circumstances. For CAS Practice exams, it may be reviewed with individual examinees or groups of examinees during post-administration exam reviews.

INSIGHTS®

The Examinee Performance Report is also made available to examinees via INSIGHTS, an interactive dashboard that allows them to download their score reports directly and explore their performance by score category and at the question level with greater flexibility. Examinees can access INSIGHTS through the MyNBME® Examinee Portal. When results become available, examinees receive an email notification with instructions for accessing INSIGHTS.