

AS SEEN IN DEANS DIGEST

Demystifying the GMAT: Reliability

By Lawrence M. Rudner

A test's reliability is the extent to which test scores are consistent over repeated sittings. This concept is critically important in standardized testing, and reliability is a key reason why test scores have meaning. Candidates with the same ability should always get close to the same score. If there is a great deal of inconsistency in the scores, then the test cannot provide a consistent assessment of what it is supposed to measure.

Here are five issues that can reduce test reliability and what we do on the GMAT exam to minimize these threats.

- 1. Administration consistency.** The entire concept of a standardized test is that the test is administered and scored in a consistent manner. Giving a test under different conditions introduces extraneous variables with unknown effects. Every GMAT exam is administered with the same software, same interface, same hardware, same number of test questions, same mix of test questions, and the same environment.
- 2. Content mix.** A test is defined by the array of skills that it measures. If the content mix were to vary, then the skills being assessed would vary, and reliability would be reduced. Every GMAT test taker will see the same number of each type of question. For the Quantitative section, the mix of data sufficiency, problem solving, algebra, geometry, arithmetic function, applied and formula-based questions will always be the same. Some computer adaptive testing algorithms let the mix vary across test takers and assure only that the average over a large number of test takers is the same. Such an approach is not consistent with our core value that the test must be fair across all examinees.
- 3. Question quality.** For a test to be reliable, the responses to the test questions must reflect the test takers' ability with regard to the intended skills. If not, then extraneous variation is introduced, and reliability is reduced. Every GMAT question is written by a professional item writer, following a detailed item-writer's guide, and every question goes through multiple rigorous reviews and statistical analyses. Only high-quality questions that are fair to all individuals and do an excellent job of measuring the intended skill survive and become eligible for inclusion in the GMAT exam.
- 4. Scale stability.** This refers to whether the scaled scores have the same meaning over time. A GMAT Total scaled score of 620, for example, means the same thing regardless of when the test was taken. Even though the pool of test questions has changed over time, all scores are mapped to the same scale that we have been using since the 1991. Because the population of GMAT test takers changes, we compute new norm tables each year to help schools and test takers compare individual candidates against the entire GMAT test-taking population.
- 5. Making sure test takers know what is expected.** Candidates, schools, and GMAC share a desire for everyone's GMAT scores to accurately reflect their abilities. Test takers need to properly prepare for the GMAT exam by becoming familiar with the item types and learning how to pace themselves. GMAC provides free practice examinations and publishes books and on-line diagnostic tools. We actively reach out to the test prep industry to provide accurate, useful information to test takers about the GMAT exam.

While there are several quality procedures that should be followed to ensure the highest level of reliability, no test is perfectly reliable. Reliability is a function of both the test and the test taker, and human beings are not perfectly reliable, either. Some small random variation of scores is always expected. As such, we encourage schools to take the standard error of measurement into account when reviewing scores. Candidates whose scores are within 30 points of each other should be treated equally.

Lawrence M. Rudner, PhD, MBA, is vice president of research and development and chief psychometrician at the Graduate Management Admission Council. He can be reached at lrudner@gmac.com.

This article was published in the August 2010 issue of *Deans Digest*. Go to <http://www.gmac.com/gmac/NewsandEvents/Deans-Digest/2010/August2010/Demystifying-the-GMAT-Reliability.htm>

© 2011 Graduate Management Admission Council. All rights reserved. This article may be reproduced in its entirety without edits with attribution to the Graduate Management Admission Council.

GMAC®, GMAC logo, GMAT®, GMAT logo, and Graduate Management Admission Council® are registered trademarks of the Graduate Management Admission Council in the United States and other countries. www.gmac.com/trademarks