

## **New Standards for New Item Types**

by Wallace Judd, Ph.D.

Classical and IRT statistics did a reasonable job of describing multiple choice questions selected to conform to normal distributions. But how do you handle items exhibiting true discontinuities? Time and accuracy are two different things – how can they be appropriately integrated into a single score? How does one create and verify a cutscore that is so complicated only a computer can calculate it? What are reasonable QA standards for a simulation? And what is a reasonable measure of reliability for a test which is never given twice in the same form?

Dr. Judd is a psychometrician with background and experience to develop illuminating answers to these and other questions in testing. After getting degrees from Princeton and Harvard, he obtained his Ph.D. from Stanford University under Lee Cronbach in IRT. A former school teacher, textbook author and programmer, he's authored numerous tests and now consults with a wide variety of companies on test design, development and evaluation. He is an ANSI examiner and President of the Board of Directors of the Performance Testing Council, a national non-profit devoted to advancing Performance Testing methodologies.