Using Response Time to Increase the Construct Validity of Trait Estimates

James A. Bovaird, University of Nebraska-Lincoln
Susan Embretson, Georgia Institute of Technology

Abstract
Response time can be used to investigate examinee cognitive processing, specifically the existence of different response strategies. In three studies, response time data from a test of abstract reasoning administered under distinct time conditions were used to increase the construct validity of an ability estimate. Study 1 used a latent profile analysis of response time to detect multiple examinee strategies to investigate the construct representation aspect of construct validity. Study 2 used mixture distribution modeling to determine the proportion of examinees that exhibited rapid-guessing behavior on each item. Item characteristics were used to account for individual item differences. Study 3 addressed the nomothetic span aspect of construct validity and showed that excluding responses made by rapid-guessers improved model fit and explained more of the variance in item performance.

Introduction
Response time (RT) has historically played a very consistent and extremely important role in testing, but that role is changing with the advent of new technologies. Due to the increased prevalence of computerized testing, response time has increased relevance to the information provided by tests. Hornbook (2000) suggested that the gain involved in using testing, simply obtaining an estimate of ability for the examinee does not utilize all of the available information. Collection of response time information on an item level can now be unobtrusive, leading to the argument that researchers have a responsibility to further investigate latency information in the pursuit of test equity and fairness (Schnipke & Scrams, 2002).

Current Psychometric Research
Current psychometric research involving response time can be categorized into several distinct areas of research (Schimdt & Scrams, 2002): the development of scoring models that either incorporate response time into the estimation of ability or model response directly as a measure of some attribute that has not been operationally defined (Cronbach & Meehl, 1955). The Common-guessing model was the best representation of the data. Removing error due to rapid guessing increased difficulty & item-level performance on the complete sample and with rapid-guessing examinees from item calibration.

Behavior Types
- "Solution" behavior (Schimdt & Scrams, 1997)
- "Rapid-guessing" behavior (Schimdt & Scrams, 1997)
- "Incorrect" behavior (Schimdt & Scrams, 1997)
- "Correct" behavior (Schimdt & Scrams, 1997)

Participants and Measures
- Abstract Reasoning Task (ART; Embretson, 1996)
- n = 818 military recruits
- Multiple item characteristics: condition (power, 70%, 50%), item difficulty, memory load

Methods
- Item-level analysis
- Comparative analysis
- Latent profile analysis
- Mixture distribution modeling

Results
- Study 2: Using the Common-guessing model, the best represent of the data, the proportion of guessing responses in response time condition.

Conclusions
- At least 2 latent classes: solution behavior & guessing behavior
- The Common-guessing model was the best representation of the data
- Proportion of guessing responses in response time condition
- Serial position was the strongest predictor of guessing behavior

References

For more information about this study, please contact James A. Bovaird
(bovaird2@unl.edu)