

# Modern Testing Needs and Student Profile Scoring

Louis V. DiBello

Educational Testing Service, Rosedale Road MS 14E, Princeton, NJ 08541

ldibello@ets.org

**Keywords:** Diagnostic classification, Student profile scoring, Bayesian inference networks, Unified model, Model based

## Abstract

Educational Testing Service has pioneered the application to large scale testing programs of new psychometric methods for student profile scoring. The College Board's Score Report Plus™ provides skill information to students who take the PSAT/NMSQT™. These psychometric methods are being applied now to a national end-of-course algebra test, to PRAXIS tests, and to the NAEP. We focus in this paper on issues that occur when an existing test is retrofitted with diagnostic scoring, and answer the question of when that is useful and reasonable. Real data examples will be provided to demonstrate strength of diagnostic performance from existing tests. We describe implications for designing diagnostic tests with strong diagnostic power.

A technical discussion will describe various modeling approaches, including the unified model cast in a Bayesian framework with MCMC estimation method called the Fusion model (Hartz) and Bayesian inference networks: Recent theoretical advances will be discussed, including equating at the skill level (Roussos, Bolt), comparisons with sub-scoring approaches, estimation of classification reliability (Douglas and others), and various approaches to test form design and computer adaptive algorithms.

## References

- Bolt D, Li Y (2003) *A low-dimensional IRT approach to linking calibrations based on the fusion model*. Technical report.
- Hartz S (2002). *A Bayesian framework for the unified model for assessing cognitive abilities: blending theory with practicality*. Ph.D. Thesis, University of Illinois.
- Hartz S, Roussos L, Stout W (2003) *Fusion model assessment of the capacity of the PSAT/NMSQT™ to diagnose skills*. Manuscript in preparation.
- Junker B, Sijtsma K (2000) *Cognitive assessment models with few assumptions and connections with nonparametric IRT*. National Research Council.
- Roussos L, Xu X, Stout W (2003) *Equating with the fusion model using item parameter invariance*. Technical report.
- Xu X, Chang H, Douglas J (2003) *A simulation study to compare CAT strategies for cognitive diagnosis*. Technical report.
- Yan D, Almond R, Mislevy R (2003) *Empirical comparisons of cognitive diagnosis models*. Technical report.