

A Paradigm for Multinational Adaptive Testing

Anthony R. Zara, 10.3.2011

Presented as part of *What's Next? A Discussion of Research Issues and Practical Requirements that May Influence the Next Generation of Adaptive Tests (invited symposium)*

Taking Your Testing Program International

- Program goals and purpose
- Localization issues
- Examination logistics
- Distance issues
- Evaluation
- Discussion

Purpose/Goals of International Testing

- Mobility of professionals
- Increasing membership
- Increasing testing volume
- Improving quality of practice
- US or local credential
- Evangelizing organization messages

Issues in “Normal” Adaptive Testing

- IRT Model
- Pretested and Calibrated Item Pool
- Ability Estimation Algorithm
- Content Control Mechanism
- Item Selection Algorithm
- Item Exposure Control Mechanism
- Stopping Rule

Goal is valid construct measurement

Issues in “Exported” Adaptive Testing

- Professional practice differences
- Cultural differences
- Different educational systems
- Translation of items
- Communications (clarity vs. standardization)
- Testing center differences
- Technology
- Different regulatory systems

Goal is construct equivalency in the new environment

Examination Logistics

- Registration
- Payment
- Test development
- Administration
- Test security
- Data flow
- Reporting

Distance Issues

- Language
- Local vs. US culture
 - Candidates
 - Officials
 - Education
- Local regulatory groups
- Score meaning

Impact of Issues in “Exported” Adaptive Testing

- Exam content issues
- Validity issues/construct differences
- Method bias
- DIF (item performance)
- Interpretation of scores
- “what do they mean?”

Goal is construct equivalency in the new environment

Evaluation

- Success factors
- Program outcomes
 - Psychometrics
 - Security
 - Volumes
 - Financials
 - Public relations
- Communications

Research Questions

- How can we assure that the constructs being measured across the geographies are equivalent?
- Do we need to assure that the educational opportunities are “equivalent” across geographies?
- Do standard CAT content balancing constraints work?
- Computerized testing is common in the US, is there a need to re-visit the basic research on modality comparability in different geographies?

Contact Information

Anthony R. Zara, Ph.D.

Pearson VUE

tony.zara@pearson.com