Evaluating Equivalence of Assessment Translations/Adaptations

A multifaceted but simplified approach.

THE SCIENCE OF PERSONALITY®

Kevin D. Meyer, Ph.D.
International Research Consultant

SIOP Leading Edge Consortium 2009
Denver, CO

Hogan Overview

- 3 Core Assessments
  - Hogan Personality Inventory (HPI)
    - 206 True-False Items
    - Normal Workplace Behavior (Five Factor Model)
  - Hogan Development Survey (HDS)
    - 168 True-False Items
    - Derailing Characteristics
  - Motives, Values, Preferences Inventory (MVPI)
    - 200 Agree-Unsure-Disagree Items
    - Values and Drivers

Equivalence Analyses

- Bootstrap Classical Test Theory Analyses
- Procrustes Analyses
- Bilingual Studies
- Adjective Checklist (ACL) Studies
- Client Validation Studies

Classical Test Theory Analyses

- Used to:
  - Flag “problematic” items/scales
  - Provide descriptive comparisons
- Single comparison samples
  - Representation
    - Generalizability of results
  - Sampling effects
    - Job type
    - Industry
    - Demographics
- Bootstrap sampling
  - Repeated sampling from single sample to create and simulate population
  - Creating parameter estimates and thresholds

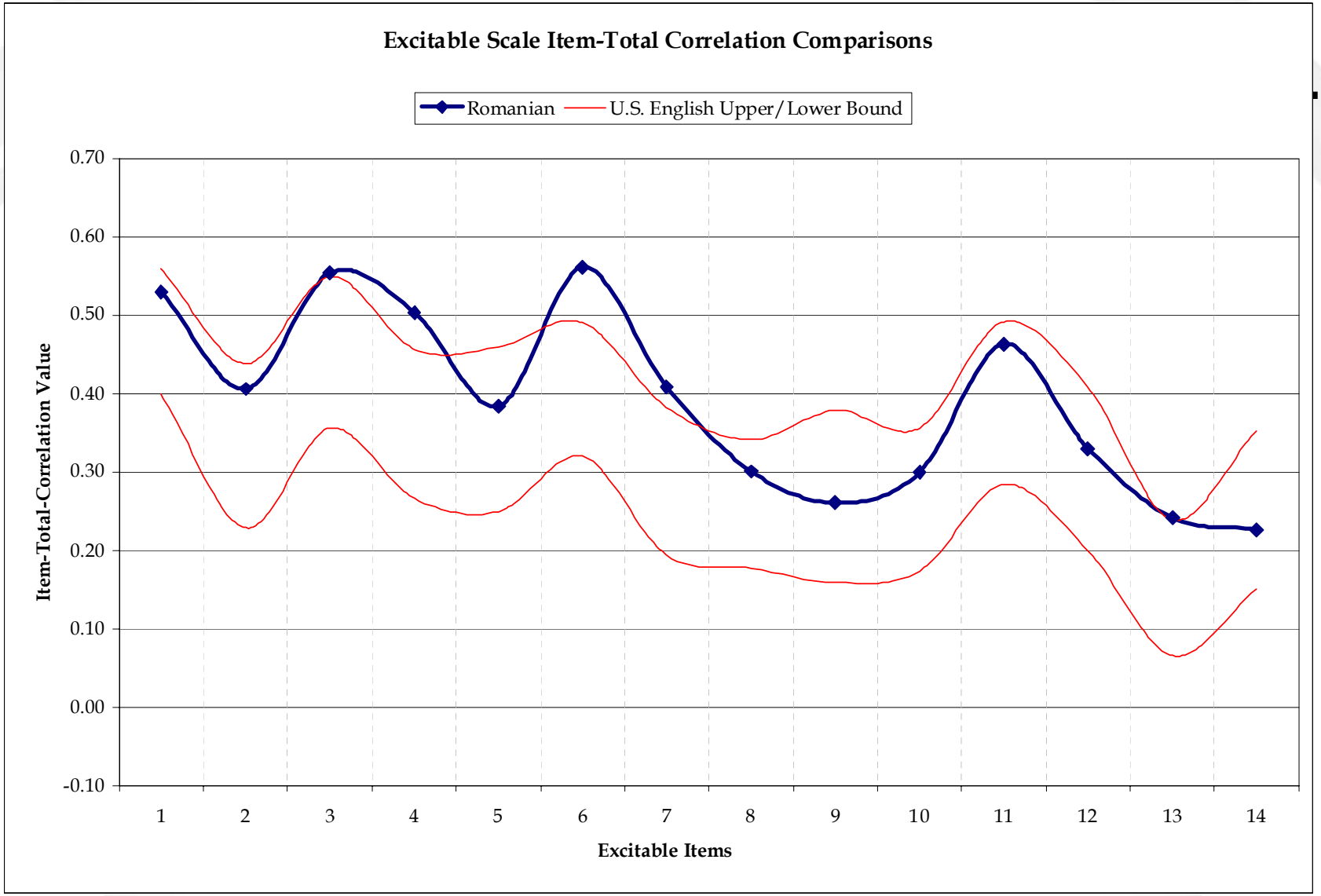
Classical Test Theory Analyses

- Comparisons between…
  1. Initial 500+ sample from new translation
     - "Convenience samples"
       - current clients
       - prospective clients
       - university samples
       - workshop samples
  2. Bootstrap sample of U.S. English data
     - Matched on sample size (1,000 iterations)
     - Can be matched on other characteristics
       - Job Type, Gender, Industry, etc.

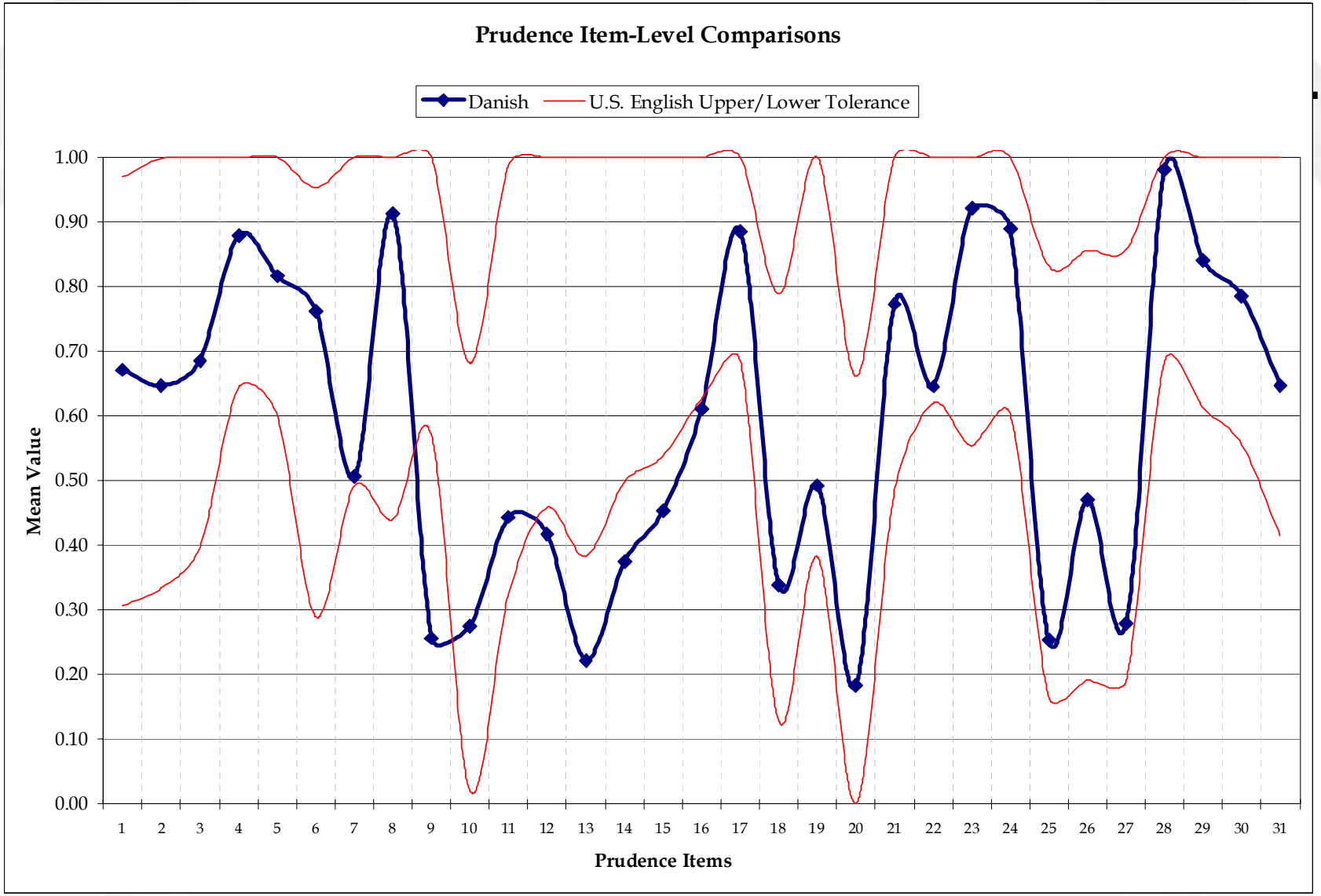
Classical Test Theory Analyses

- Bootstrap sampling creates distributions of…
  - Scale level statistics
    - Mean
    - Internal consistency (Cronbach's alpha)
  - Item level statistics
    - Mean (i.e., endorsement rate)
    - Corrected item-total correlations
- Construct thresholds for acceptable statistics in non-English sample
  - 2.5^{th} and 97.5^{th} percentiles

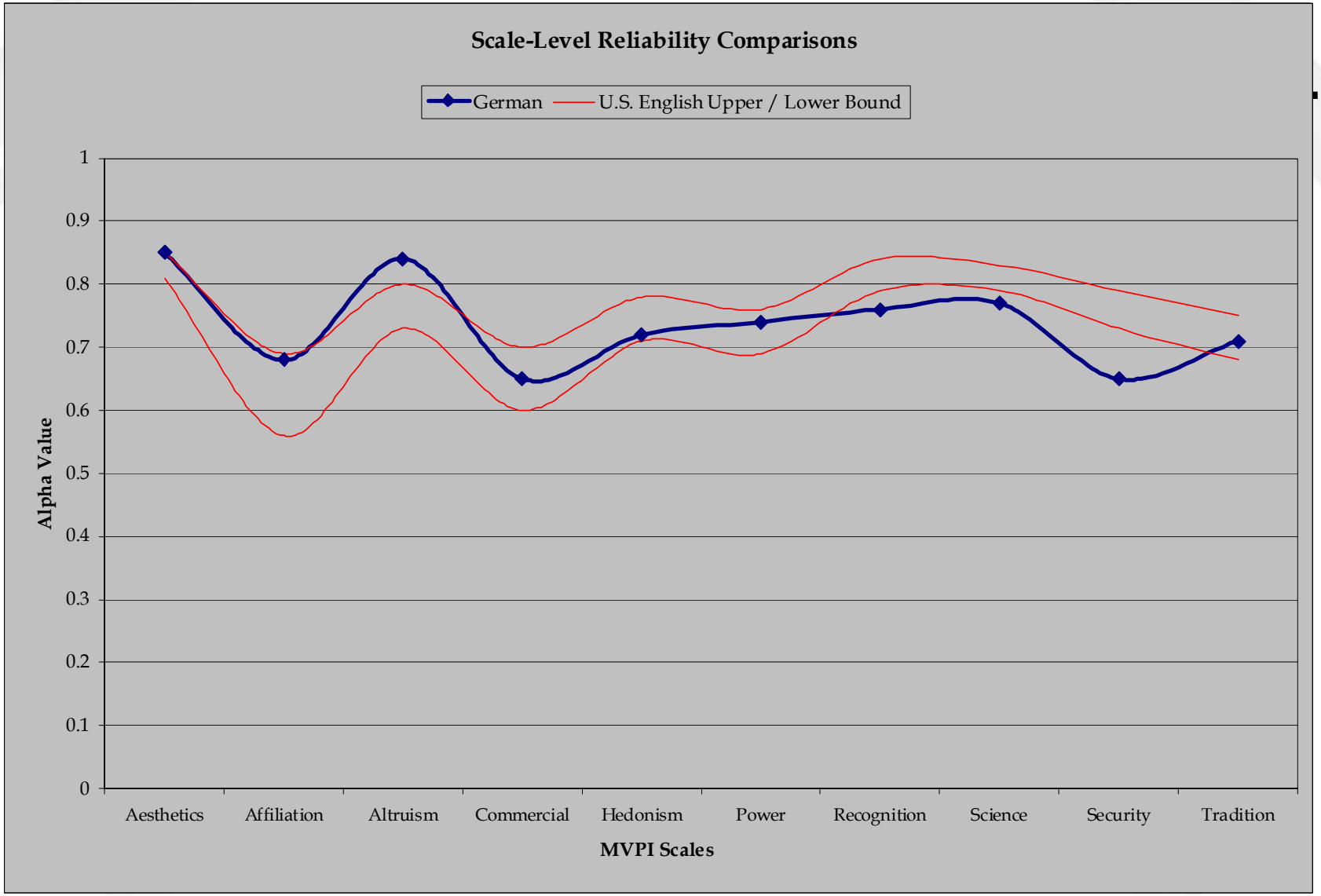
Classical Test Theory Analyses



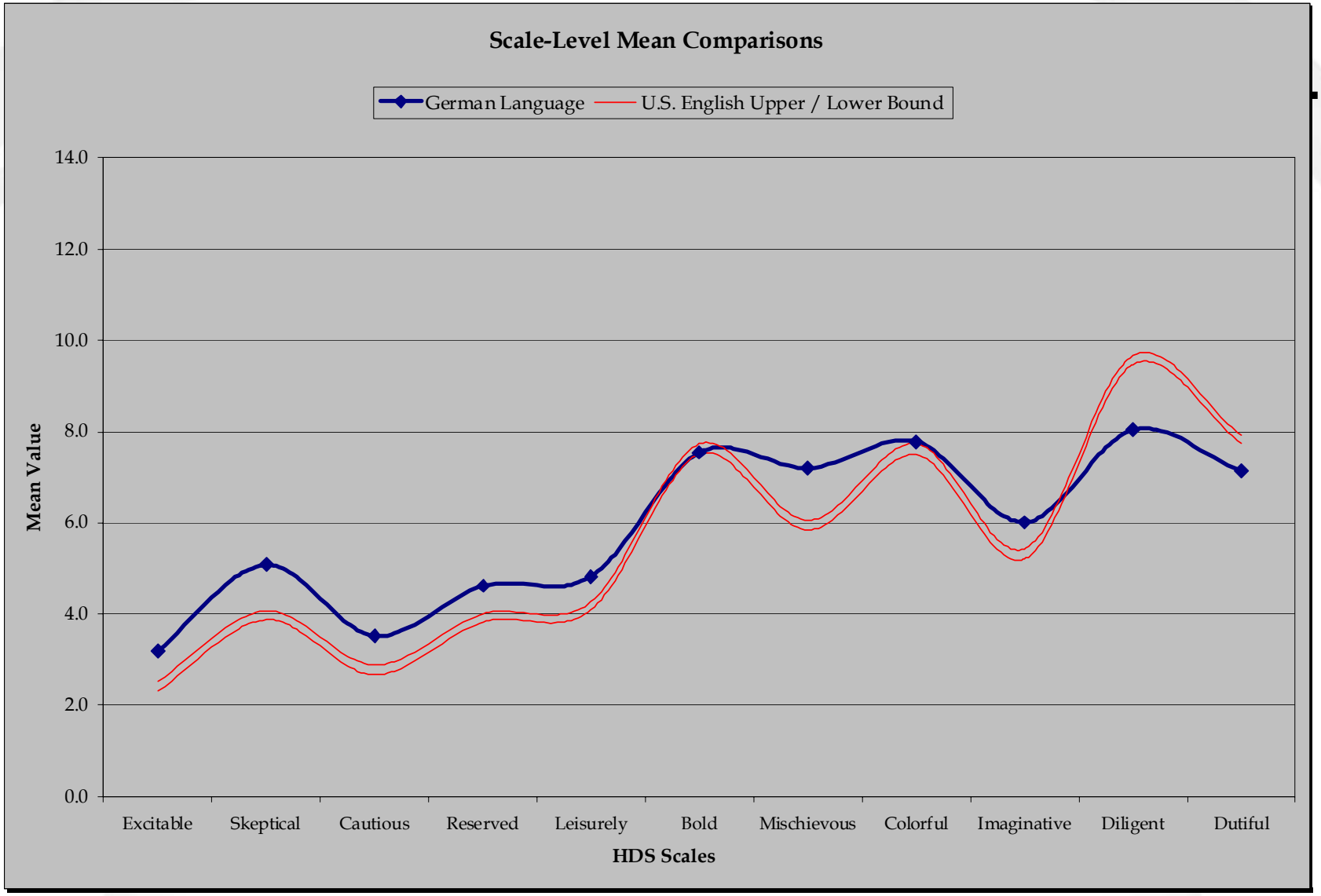
Classical Test Theory Analyses



Classical Test Theory Analyses



Classical Test Theory Analyses



Classical Test Theory Analyses

- What do you need to do this?
  - Data from translated form (e.g., >500 cases)
  - Large comparison dataset (e.g., US English normative dataset)
  - SPSS
    - Patience and savvy with syntax creation!
  - A strong PC

Procrustes Analyses

- Evaluates congruence of factor structure
- Compares factor loading matrices
  - Spanish factor loading matrix is Procrustes rotated to target US English factor loading matrix
- Analyses yield congruence coefficients (-1 to 1)
  - Assessment, factor/scale, and subscale
  - >.85 (or .90) considered similar

Procrustes Analyses

- Eval … re
- Com …
  - Sp … es
    ro … g matrix
- Anal … (-1 to 1)
  - As …
  - >. …

FACTOR CONGRUENCE COEFFICIENTS

	MOVING A	MOVING A	MOVING T	ITEMCONG
EXC	.723	.004	-.250	.975
SKE	.483	.267	.186	.976
CAU	.706	-.387	.106	.997
RES	.585	-.128	-.036	.989
LEI	.519	.186	.175	.999
BOL	.129	.655	.252	.987
MIS	.050	.645	-.110	.995
COL	-.247	.677	-.258	.978
IMA	.086	.602	-.082	.983
DIL	.038	.051	.763	1.000
DUT	.169	-.122	.233	.916
FACTCONG	.989	.992	.948	.981

------ END MATRIX -----

Procrustes Analyses

- What do you need to do this?
  - Data from translated form (e.g., >500 cases)
  - Comparison dataset (e.g., US English normative dataset)
  - SPSS
    - Patience and savvy with syntax creation!

Bilingual Studies

- Bilingual participants complete English and translated form of assessment
  - 30-50+ participants
  - Balanced design
- Eliminates people effects
  - Within subjects design
- Statistics
  - D, |D|, correlations, agreement indices (ICC)
- Compared to English-English test-retest sample
  - Similar interval (e.g., <7 days)

Bilingual Studies

- What do you need to do this?
  - Test-retest sample from original form
  - A bilingual sample!
    - Universities

Adjective Checklist Studies

- Getting back to construct validity
- Participants complete assessment
- Participants nominate 1-2 others to complete ACL
  - Rate participant on list of descriptive adjectives / adjectival statements
    - E.g., “Appears to be self critical”
- Do observer descriptions (ACL responses) align with assessment scores?

Adjective Checklist Studies

- What do you need to do this?
  - An Adjective Checklist…translated!
  - A sample
    - Universities

Client Validity Studies

- Getting to the heart of the matter
- Local evidence of predicting job performance
- Predictive / Concurrent Criterion-Related Validation Studies
- Do the same characteristics predict performance in the non-US job as we see in the US?
  - Yes = evidence of job similarity and transportability of validity evidence.
  - No = different characteristics predicting success? Cultural difference?
  - No = nothing is predicting success? Uh oh.

Client Validity Studies

- What do you need to do this?
  - A receptive client!

Summary

There is more than one way to evaluate equivalence.

More "sophisticated" approaches may not be necessary.

Accept that perfect measurement equivalence will not be achieved.