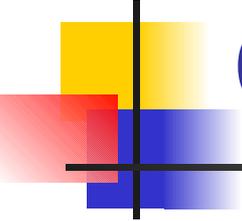# Comparability of paper and computerized non-cognitive measures: A review and integration
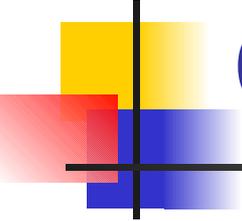
Alan D. Mead

American Institute of Certified Public Accountants

David L. Blitz

Chicago School of Professional Psychology and Illinois Institute of Technology

# Comparability is…

- Different versions (paper- and computer-based) produce the same scores
  - Same rank order of test-takers
  - Same mean and standard deviation
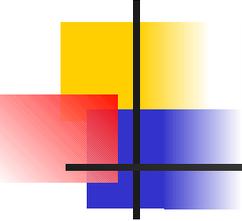  - Same correlation with other variables

# Comparability Studies

- Between Subjects
  - ANOVA/ANCOVA/etc.
  - DIF
  - SEM
- Within-Subjects
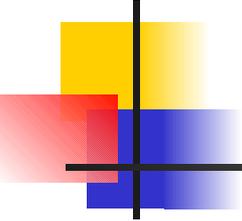  - Correlation
  - Paired t-test
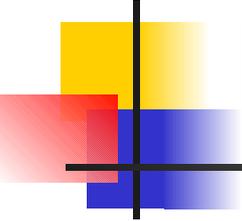
# Between Ss designs

- ANOVA/ANCOVA/etc.
  - Tests difference in means, not rank-order
  - Beware that standard statistical tests are of limited value… comparability=accepting the null
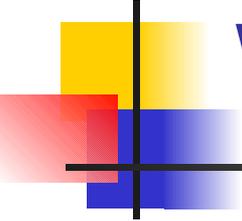  - Readily meta-analyzable

# Between Ss designs

- Differential Item Functioning (DIF)
  - Tests difference in item means but not rank-order nor group differences
  - Probably more powerful than ANOVA
  - Again, standard statistical tests are of limited value
  - Cannot be cumulated using meta-analysis

# Between Ss designs

- Structural Equations Modeling (SEM)
    - Tests equality of correlations
    - Also, tests difference in item and/or scale means but not rank-order
    - Probably more powerful than ANOVA
    - Standard statistical tests do not directly test equivalence (you accept the null)
    - Cannot be cumulated using meta-analysis

# Within Ss designs

- Correlation
  - Directly tests rank orders, but does not test means
  - Common statistical tests not valuable
  - Readily meta-analyzable
- Paired t-test/repeated measures
  - Test means, more powerful than ANOVA
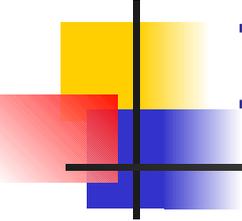
# Ideal study

- Large Within-Ss sample
  - Assess correlation
  - Use equivalence testing (Rogers, et al., 1993) to assess means
  - Use SEM to analyze covariance
- Also, include covariates
  - Include measures that might explain non-comparability
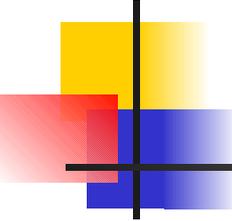
# Meta-analysis

- Found 105 studies comparing non-cognitive paper- and computer-based versions
  - Only 6 studies were appropriate
  - within-Ss, noncognitive measures, with reliabilities
  - N=760, 41 correlations
  - Weighted average cross-mode correlation=1.02