

Testing Then and Now:
Building on a Legacy in Educational
Measurement, Assessment and Evaluation

Some Milestones in the History of Educational
Measurement

Robert L Brennan
University of Iowa

A Presentation at Teachers College, Columbia University
December 9, 2013

Organization of Comments

- 1900-1950: The Beginnings; Some Contributions of E. L. Thorndike
- 1950-Present: Vigorous and Varied Development and Use of Measurement Models
- Enduring Issues
- The Future

Edward Lee Thorndike (1874-1949)

“Anything that exists,
exists in some amount and can be measured”

- Student of William James at Harvard (1896-1897)
- Student of Cattell at Columbia (*Animal Intelligence*, 1898)
- Correlation Lineage: (Galton -> Pearson) ->Cattell -> Thorndike
- Began academic career at Columbia in 1899
- Father of “modern” educational psychology
- Productivity:
 - 508 publications in nearly every major area in psychology and education (e.g., intelligence testing, achievement testing, interest inventories, educational psychology, statistics, etc.)
 - “the most productive psychologist our country has produced” (Woodworth)
 - Influence on CTT (e.g., Army Alpha and Beta tests; paper on equating (1922); advisor of Truman Kelley; normalized T-scores named after Thorndike and Terman)
- Disliked “any abstract discussion not tied to concrete facts (Woodworth)”
- “tolerant of divergent views though always asking for ‘the evidence’ “ (Woodworth)
- E.L. Thorndike facilitated the development of sound theory and practice in testing -
-- sometimes directly; sometimes indirectly

The Beginnings: 1900-1950

Classical Test Theory and Factor Analysis

- Classical Test Theory Model: $X = T + E$
- Spearman (2004): correction for attenuation (brilliant insight)
- Spearman-Brown Formula (2010)
- Kuder and Richardson (1937) (KR-20 & KR-21)
- Hoyt (1941): used ANOVA to estimate reliability
- Gulliksen (1950) --- still useful and relevant
- Cronbach's coefficient alpha (1951)
 - Most used AND misused estimate of reliability
 - Somewhat of an embarrassment for Cronbach
 - Precursor to generalizability theory
- Factor analysis (Spearman, Thurston, many others)
- Virtually all of validity was predictive validity
- NOTE: Influence of World War's I and II

1950-Present: Vigorous and Varied Development and Use of Measurement Models

- Models
 - Lord and Novick (1968) *Statistical Theories of Mental Test Scores*
 - *Generalizability Theory* (Cronbach et al., 1972, Brennan, others)
 - *Item Response Theory* (Lord, Rasch, Bock, Mislevy, others)
 - *Equating and Scaling* (Angoff, others)
- Five (soon to be six) editions of *Standards for Educational and Psychological Testing*
- Four editions of *Educational Measurement* (edited by Lindquist, R. L. Thorndike, Linn, and Brennan, respectively)
- Numerous developments in validity/validation (Cronbach, Messick, Kane)
- Vastly increased amount of educational testing
- Increasing use of computers in various aspects of testing
- The politics of testing have taken on ever increasing importance

Enduring Issues/Tensions

- Ability vs. Achievement Testing
- Norm-referenced vs. Criterion-referenced
Interpretations
- Formative vs. Summative Evaluation
- Measuring Status vs. *Growth*

The Future

- Need for
 - **integration** of measurement models
 - use of improved **validation** practices (e.g., Kane)
 - adoption of richer assessment frameworks and formats that capitalize on APPROPRIATE use of **technology**
 - improved **communication** of assessment results to different audiences
 - many more trained **measurement professionals**
- I am optimistic, but the challenges are formidable

Thank you for the opportunity to be
part of this celebration of the
contribution of Teachers College to
educational measurement.