UCLA Semel Institute
Tennenbaum Center for the Biology of Creativity

UFHealth
UNIVERSITY OF FLORIDA HEALTH

MEDICAL COLLEGE OF WISCONSIN

EMORY UNIVERSITY

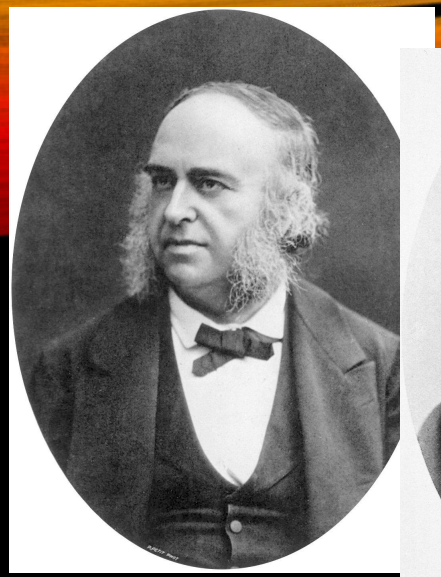
TOWARD PRECISION NEUROPSYCHOLOGY

Modern Psychometric Strategies for Precision Neuropsychology

Laura Glass Umfleet, PsyD, ABPP(CN)

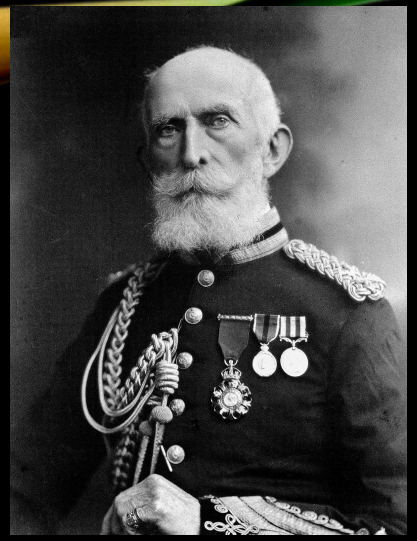
OUTLINE

- Phases of neuropsychology (Neuropsychology 1.0 and 2.0)
- Psychometric theory & modern psychometric strategies
- Neuropsychology 3.0



This Photo by Unknown Author is licensed under CC BY-SA

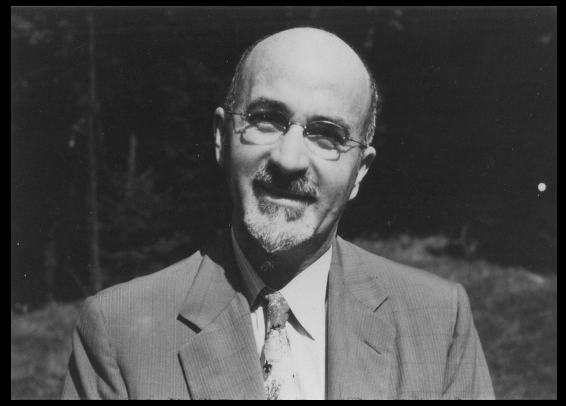
This Photo by Unknown Author is licensed under CC BY-SA



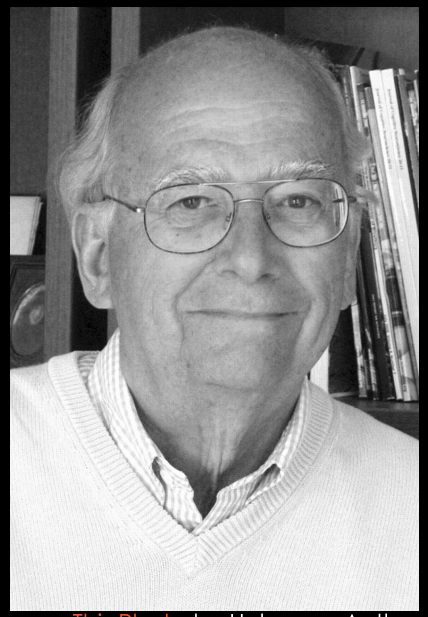
This Photo by Unknown Author is licensed under CC BY-SA

Photo by Unknown Author is licensed
er CC BY-SA

This Photo by Unknown Author is licensed under CC BY-SA



This Photo by Unknown Author is licensed under CC BY



This Photo by Unknown Author is licensed under CC BY

PHASES OF NEUROPSYCHOLOGY

- Neuropsychology 1.0 (1950-1979)
  - Neuropsychology recognized as a discipline in the 1960s
  - Neuropsychologists typically practiced in neurology clinics and investigated functions associated with specific brain lesions
  - Interpretation based on assessment without adequate normative data
  - Development of some formalized batteries
- Neuropsychology 2.0 (1980-present)
  - Widespread availability of neuroimaging
  - Classical psychometrics with newer tests improving on standardization and co-norming
  - Growth of symptom validity testing

Bilder (2011)

NEUROPSYCHOLOGY 2.0

- Most frequently used NP tests have roots in the 19th century
- Timely and costly
- Lack integration with current technologies

ASSESSMENT AIN'T EASY

- Test items are complicated devices and people are even more complex
  - Different people probably have different "wiring" and there is much to learn about the complexities of the brain
  - No two people possess the same knowledge base and have the same thought processes when answering test items
- Measurement of skills or knowledge in neuropsychology has largely been based on Classical Test Theory

CLASSICAL TEST THEORY

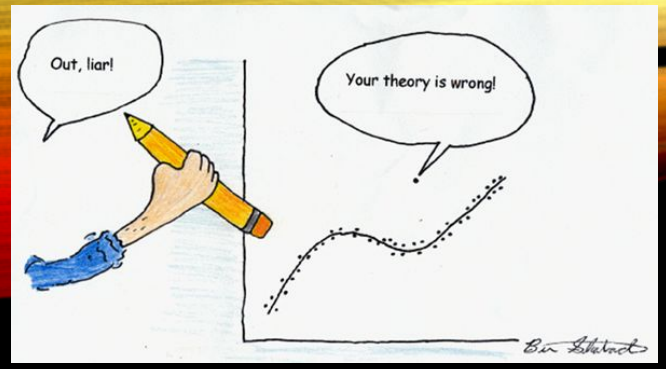
- Aim: understand and improve the reliability of psychological tests

CLASSICAL TEST THEORY

- **Classical test theory**- X = T + E
  - **True score**= the score an examinee would obtain on a measure in the absence of error
  - **Error component**= measurement of error
  - **Measurement of Error**= random error due to factors that are irrelevant to what is being measured and have an unpredictable effect on test score

CLASSICAL TEST THEORY

- **An obtained test score reflects truth and error**
  - Considers item difficulty and discrimination, reliability, and validity
- **Weaknesses**
  - Test items and parameters depend on the sample tested
  - Difficult to equate scores on different tests



IRT

- Popularized by the work of Fredrick Lord and Georg Rasch in the 1950s and 1960s
- Latent trait(s) gives rise to an individual's responses to individual test items
- Theoretically, item characteristics are not sample dependent

	CTT	IRT
Model	Linear	Non linear
Level	Test	Item
Score	Depends on items	Item independent
Item parameters	Sample dependent	Sample independent
Preferred items	Average difficulty	Any difficulty

Reckase, 2009; Reise & Waller, 2009

IRT

- Unidimensional models
  - Item-level responses are analyzed to compare the probability of a correct answer against the underlying trait or ability level
  - Most IRT analyses have assumed a single latent trait underlies responses
- Bifactor models
  - Items load on both a general dimension and individual factors
- Multidimensional models
  - Each item can provide information about multiple different traits

Reckase, 2009; Reise & Waller, 2009

IRT

- IRT has advantages over classical test theory for neuropsych assessment
  - Nominal response model
  - Test linking
  - Computerized adaptive testing
  - Differential item functioning
  - Application of person-fit statistics

Reise & Waller, 2009

IRT

- Nominal response model
  - Tests with more than one response option (e.g., multiple choice tests)
  - Possible unique information from wrong responses
  - May be valuable with increasingly accessibility of online tests
- Test linking
  - Items from different tests can be placed on a common scale
  - Use of IRT-based methods would inform if tests of the same ability tap the same latent trait
  - Provides opportunity for development of new measures

IRT

- Computerized adaptive testing
  - Enhance efficiency of neuropsych testing
    - Efficiency gains of 50 to 95% without negatively affecting quality
  - Select most informative items on given trials to update estimate of an examinee's ability
  - Can efficiently examine multidimensional or bifactor IRT models

- Differential item functioning
  - A given item may behave differently in different groups with the same true ability
  - Potential utility in evaluation of persons from different cultural and linguistic backgrounds
  - There are multiple IRT-based approaches to examine differential item functioning

IRT

- Person fit statistics
  - Identify abnormal patterns (e.g., guessing behaviors) of item responses that do not fit with identified trait for an individual or general patterns observed in other examinees
  - Potential to help with development of imbedded performance validity tests

DISADVANTAGE OF IRT

- Need large sample size and need several scale items to estimate latent trait scores and item parameters.
- Can be cumbersome when measurement invariance needs to be evaluated across multiple groups
- IRT is good for application for many tests but not for all
  - Some tests can be modified to apply IRT models
  - If not, leveraging technology to collect item level data is possible

e.g., Meade & Lautenschlager, 2004

NEUROPSYCHOLOGY 3.0

- Ontology development
- Create data repositories
- Increase use of technology

Bilder (2011)

LEVERAGING TECHNOLOGY

- Current NP practice: mostly rely on print publishing of various measures, manual calculation of scores, look up norms in manuals, and then enter scores on a data summary sheet.
- Future:
  - Increase use of computerized versions of published tests
  - Further growth of web-based testing
  - Healthcare and bio informatics
  - Mobile platforms
  - Wearables

REFERENCES

Bilder, R. M. (2011). Neuropsychology 3.0: Evidence-based science and practice. *JINS, 17,* 7-13.

Bilder, R.M., Howe, A. S., Sabb, F. W., & Parker, D. S. (2013). Multilevel models from biology to psychology: Mission impossible? *Journal of Abnormal Psychology*.

Bilder, R. M., & Reise, S. P. (2019). Neuropsychological tests of the future: How do we get there from here? *TCN, 33,* 220-245.

Choi, S. W., Reise, S. P., Pilkonis, P. A., Hays, R. D., & Cella, D. (2010). Efficiency of static and computer adaptive short forms compared to full-length measures of depressive symptoms. Quality of Life Research, 19(1), 125–136.

Reckase, M.D. (2009). Multidimensional item response theory. Springer, NY.

Reise, S. P., & Waller, N. (2009). Item response theory and clinical measurement. Annual Review of Clinical Psychology, 5, 27-48.

Waller, N. G., & Reise, S. P. (1989). Computerized adaptive personality assessment: An illustration with the absorption scale. Journal of Personality and Social Psychology, 57(6), 1051–1058.