

Test Theory  
PSYCHOLOGY 842  
FALL 2022

**Location:** T 9-11:30am at Davie Hall 347

**Instructor:** Oscar Gonzalez

**Office:** 337A Davie Hall (3<sup>rd</sup> floor)

**Office Hours:** By appointment

**Email:** [ogonza13@unc.edu](mailto:ogonza13@unc.edu)

**Website:** Sakai class site

**Prerequisites:** PSYC830

### Reference Text

Bandalos, D. L. (2018). *Measurement theory and applications for the social sciences*. Guilford Publications. [referred to as **DB**] \*\*

\*\*I recommend getting the book for your own reference, but copies of chapters will be provided in Sakai. Supplemental readings will also be assigned and posted.

### Statistical Software (FREE)

In this class we will use the R statistical language for data analysis. R is a popular, script-based statistical language that you can download for free in your own laptop.

Download: <https://www.r-project.org/>

### Overview and Learning Goals for the Course

In this day and age, assessment is an integral part of many institutions and disciplines. From assessing educational achievement, health outcomes, or psychological constructs, we use scales and measures to make inferences and decisions about individuals. Psychometrics can be defined as the study of the measurement of psychological constructs, which ultimately “encourages the development of psychology as a rational science” (Thurstone, 1937). Test theorists develop scales and evaluate the inferences we can make from them. This course is an introductory seminar on measurement and psychometric theory.

The overall goals of the course are to...

1. Understand how to evaluate constructs precisely
2. Understand how to ascribe a valid interpretation to the score on a measure
3. Understand the role of factor analysis and item response theory in scale construction

One of the major aims of this course is for students to think critically about measurement in their own research, so the target audience is psychology graduate students with a basic stats background.

### Structure

The general structure of the course is the following. Most days I will lecture on a measurement topic, discuss its theoretical underpinnings, and demonstrate how to conduct some analyses. Time-permitting, we will discuss questions from the readings generated by the students.

## Course Requirements, Grading, and Honor Code

I do not tolerate cheating. The honor code is in effect for this course. Students shall not misrepresent others' work as their own and will give full credit for others' contributions to the extent that these are allowed within the parameters of the assignment. Individual assignments are to be conducted individually.

Course grades will be based on three major components:

- **Participation (20%):**
  - **In-class activities (10%):** There will be several in-class activities, discussions, and online forum interactions. Regular class attendance is expected.
  - **Discussion questions (10% - Starting Week 3):** Everybody is to read the assigned articles/chapters prior to class and write one (or optionally more) questions that can be the focus of clarifying discussion during class. Those questions are to be emailed to me (ogonzal3@unc.edu) by 12pm on each Monday preceding class. During class, we will (jointly) do our best to deal with the questions, either through discussion or additional reading material.
- **Three assignments (30%):**
  - Classical test theory
  - Validity theory
  - Factor analysis
- **Final Project (50%):** The goal of the project is to apply some of the skills learned in the class to your research, present it in a conference-style talk, and write a paper. Potential paper types are a substantive application or a simulation study. The final project has 3 components:
  - **Flash talk (5%):** 7min presentation where you tell the class what question you will be answering in your final project, the data that you will use, and the proposed analysis.
    - Record your talks and upload
    - Presentation viewing date is on 10/18
    - Classmates will provide feedback on Sakai Forums + presenter will respond
  - **Final presentation (15%):** 15-20 min presentation about your findings.
    - Presentations will be recorded
    - Presentation viewing date is on 11/22, offline
    - Classmates will provide feedback and questions in Sakai forums + presenter will respond
- **Paper (30%):** 12pg (double-spaced) write-up of your findings.
  - APA style: Intro-method-results-discussion, 12ft Times New Roman

The grade scale follows:

H	90-100% highest level of attainment
P	70-89%
L	60-69%
F	59 or below%

## Tentative Schedule\*

Date	Week	Lecture	Reference Reading
08/16	1 <sup>#</sup>	Introduction to Psychometrics	<b>DB1</b> ; Wijsen & Borsboom (2021) Jones (1971)
08/23	2	Historical and Statistical Background	Wintroub (2020); Longreads Post Jones & Thissen (2007)
08/30	3 <sup>#</sup>	Classical Test Theory: Reliability	<i>Standards</i> (Ch2); <b>DB7 &amp; 8</b> Algina & Penfield (2009); Schmitt (1996)
09/06	4	WELLNESS DAY	Please rest!
09/13	5	Validity I	Cronbach(1971), Angoff (1988) Cronbach & Meehl (1955)
09/20	6 <sup>#</sup>	Validity II	Messick(1995), Kane(1992), <i>Standards</i> (Ch1) Borsboom et al. (2009), <b>DB11</b>
09/27	7	Factor Analysis	Raykov & Marcoulides (2011,Ch. 3), <b>DB12</b> MacCallum (2007), Fabrigar et al. (1999)
10/04	8 <sup>#</sup>	Confirmatory Factor Analysis	Brown (2015; Ch 2, 3), <b>DB13</b> West et al. (2012)
10/11	9	Measurement Bias and Fairness	Millsap & Olivera-Aguilar (2012), <b>DB16</b> Meredith & Teresi (2006), <i>Standards</i> (Ch3)
10/18	10 <sup>@</sup>	ZOOM Flash Talks	Good luck!
10/25	11	Item Response Theory: Background	Thissen & Wainer (2001); Edwards (2009) Thissen & Steinberg (2007), <b>DB14</b>
11/01	12	Item Response Theory: Applications	Wirth & Edwards (2007); Edwards & Edelen (2009); Reise & Waller (2009)
11/08	13	Tests for selection and decision-making	Gonzalez (2021a); Gonzalez & Pelham (2021) <b>DB17</b> , Gonzalez et al. (2021; in press)
11/15	14	Test Construction	Edelen & Reeve (2007), Reeve et al. (2007) Gonzalez (2021b), <b>DB3</b> , <i>Standards</i> (Ch4)
11/22	15	ZOOM Presentations	Good Luck!
11/29	☺	FINAL PAPERS DUE @ 12pm	You got this!

**DB** refers to a chapter in Bandalos (2018)

\* Instructor reserves the right to make changes to the syllabus, including presentation dates and due dates. These changes will be announced as early as possible.

# Assignment posted, due a week from then.

@ Project proposal presentation

## References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.  
\*[aka *The Standards*]
- Algina, J., & Penfield, R.D. (2009). Classical test theory. In R.E. Millsap and A. Maydeu-Olivares (Eds.), *The Sage Handbook of Quantitative Methods in Psychology* (pp. 93-122). Thousand Oaks, CA: Sage.
- Angoff, W.H. (1988). Validity: An evolving concept. In H. Wainer & H.I. Braun (Eds.), *Test validity* (pp. 19-32). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc
- Bandalos, D. L. (2018). *Measurement theory and applications for the social sciences*. Guilford Publications
- Borsboom, D., Cramer, A. O. J., Kievit, R. A., Scholten, A. Z., & Franić, S. (2009). The end of construct validity. In R. W. Lissitz (Ed.), *The concept of validity: Revisions, new directions, and applications* (pp. 135–170). IAP Information Age Publishing.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.
- Cronbach, L. J. (1971). Test validation. In R. Thorndike (Ed.), *Educational measurement* (2nd ed., p. 443). Washington DC: American Council on Education.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research*, 16, 5-18.
- Edwards, M. C. (2009). An introduction to item response theory using the Need for Cognition Scale. *Social and Personality Psychology Compass*, 3, 507-529.
- Edwards, M. C. & Edelen, M. O. (2009). Special Topics in item response theory. In R. Millsap & A. Maydeu-Olivares, *The Sage handbook of quantitative methods in psychology*. London: Sage Publications.
- Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272-299.
- Gonzalez, O. (2021a). Psychometric and machine learning approaches for diagnostic assessment and tests of individual classification. *Psychological Methods*.

- Gonzalez, O. (2021b). Psychometric and machine learning approaches to reduce the length of scales. *Multivariate Behavioral Research*
- Gonzalez, O., Georgeson, A. R., & Pelham, W. E. III. (in press). How accurate and consistent are score-based assessment decisions? A procedure using the linear factor model. *Assessment*. DOI: 10.1177/10731911221113568.
- Gonzalez, O., Georgeson, A. R., Pelham, W. E. III., & Fouladi, R. T. (2021). Estimating classification consistency of screening measures and quantifying the impact of measurement bias. *Psychological Assessment*, 37, 596-609.
- Gonzalez, O. & Pelham, W. E. III. (2021). When does differential item functioning matter for screening? A method for empirical evaluation. *Assessment*, 28, 446-456.
- Jones, L. V. (1971). The nature of measurement. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington, DC: American Council on Education
- Jones, L.V., & Thissen, D. (2007). A history and overview of psychometrics. In C.R. Rao and S. Sinharay (Eds.), *Handbook of Statistics 26: Psychometrics* (pp. 1-27). New York, NY: Elsevier.
- Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112(3), 527-535
- MacCallum, R.C. (2007). Factor analysis. In R.E. Millsap and A. Maydeu-Olivares (Eds.), *The Sage Handbook of Quantitative Methods in Psychology* (pp. 123-147). Thousand Oaks, CA: Sage.
- Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical care*, S69-S77.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.
- Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 380-392). New York, NY: Gilford Press.
- Raykov, T., and Marcoulides, G.A. (2011). *Introduction to psychometric theory*. New York, NY: Routledge.
- Reeve, B.B., Hays, R.D, Bjorner, J.B., Cook K.F., Crane, P.K., Teresi, J.A., Thissen, D., Revicki, D.A., Weiss, D.J., Hambleton, R.K., Liu, H., Gershon, R., Reise, S.P., & Cella, D (2007). Psychometric evaluation and calibration of health-related quality of life items

banks: Plans for the patient-reported outcome measurement information system (PROMIS). *Medical Care*, 45, S22-31.

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

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8, 350-353.

Thissen, D., & Steinberg, L. (2007). Item response theory. In R.E. Millsap and A. Maydeu-Olivares (Eds.), *The Sage Handbook of Quantitative Methods in Psychology* (pp. 148-177). Thousand Oaks, CA: Sage.

Thissen, D., & Wainer, H. (Eds.). (2001). Test scoring. Lawrence Erlbaum Associates Publishers.

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). New York, NY: Guilford

Wijisen, L. D., & Borsboom, D. (2021). Perspectives on psychometrics Interviews with 20 past psychometric society presidents. *Psychometrika*, 86(1), 327-343.

Wintroub, M. (2020). Sordid genealogies: a conjectural history of Cambridge Analytica's eugenic roots. *Humanities and Social Sciences Communications*, 7, 1-16.

Wirth, R. J., & Edwards, M. C. (2007). Item factor analysis: Current approaches and future directions. *Psychological Methods*, 12, 58–79