

**Syllabus**

**Physics Education Study Program  
Faculty of Mathematics and Science**

**Instructor** : Dr. Dadan Rosana, M.Si.  
**Office** : Phsics Education Study Program  
**Office hours** :  
**Email** : danrosana.uny@gmail.com

**Course Description:**

This course introduces classical test theory, including definitions and formulas for test reliability, standard error of measurement, and related statistics. Additional topics include test validity, item statistics useful in test construction, score scales and norms commonly used in educational testing, item bias and test bias, and ideas of fairness and equity in educational and psychological testing. Factor analysis as well as the major extensions and alternatives to classical test theory, generalizability theory and item response theory (latent trait theory), are briefly introduced.

This course is intended to equip students to read the literature in their own substantive areas more critically, to use tests more intelligently in research, and to pursue further studies in psychometrics. It is prerequisite to Education (Item Response Theory) and Education (Generalizability Theory). Although the focus of the course is on paper-and-pencil measures of cognitive abilities and academic achievement, most of the concepts and methods developed apply equally to performance testing, as well as the assessment of attitudes and personality constructs, ratings based on systematic observations, and other kinds of assessments of individuals or groups.

**Materials**

There is also a collection of "programmed instruction" materials for this course, which provide a systematic introduction to and practice in using the notation, statistical concepts, and algebraic manipulations required. Various authors, primarily L. J. Cronbach, have contributed to these materials over the years. In 1982, the teaching assistant for this course, David King, retyped and extensively revised these materials to standardize their format, correct minor errors, and provide coverage of some topics that had been omitted. Some students have found these materials quite helpful. Two copies are on reserve in YSU Library.

Students are also encouraged to pursue the readings listed in the attached bibliography. These are intended both to provide additional review and practice for students especially insecure about their statistical preparation, and to offer more thorough coverage of selected topics, for those who wish to go beyond what is provided in the text or who are concerned with specific kinds of applications of measurement theory. In previous years, optional readings have been placed on reserve. The library has found, however, that the demand for optional readings is generally low, and that the restricted loan period for materials on reserve inconveniences those students who do make use of them. For the past several years, I have followed the library's recommendation in not placing these materials on reserve, and students have not reported any difficulties. If access to or competition for these materials should pose any problem, please let me know, and I'll try to locate additional copies for you to use.

## Course Schedule

<u>Week</u>	<u>Topic</u>	<u>Readings<sup>a</sup></u>
1	Introduction and Overview	C&A 3-15; PI 1-39
2	Properties of composite scores	C&A 16-50, 60-64, 87-101; PI 40-111; WE&C 56-73, Ch. 12
3	The classical (weak) true-score model	C&A 36-42(review),105-30;PI 112-133, §§B,Q; WE&C 178-192; Gulliksen Chs. 2, 3
4	Estimating reliability (1)	C&A 131-43; PI C1-C11; Feldt & Brennan 105-113; Gulliksen Chs. 6-7
5	Estimating reliability (2); stratified alpha	Feldt & Brennan 113-118; Feldt (1990)
6	Useful applications (1)	C&A 143-56; Gulliksen Chs. 8,10
7	Optional discussion section	
8	Useful applications (2); rel. of group means	Feldt & Brennan §§ 2.7.1, 2.8.8
9	Conditional standard errors of measurement	Feldt & Brennan §§ 2.8.3
10	Optional discussion section	
11	Generalizability theory	C&A 157-91; Shavelson & Webb
12	Validity (1)	C&A 217-42; PI C11-C17; Campbell & Fiske (1959); APA <u>Standards</u>
13	Optional discussion section	
	Validity (2); <u>MIDTERM DUE</u>	Messick (1989, 1995)
14	Factor Analysis	C&A 287-308
15	Test construction and item analysis	C&A 66-86, 311-38; Haertel (1985); Millman & Greene; WE&C 196-203 Petersen, Kolen, & Hoover
16	Testing and Educational Policy	

### References

- American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME). (1999). Standards for educational and psychological testing. Washington, D.C.: American Psychological Association.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, *56*, 81-105.
- Crocker, L., and Algina, J. (1986). Introduction to Classical and Modern Test Theory. New York: CBS College Publishing.
- Cronbach, L. J. (1976). Equity in selection: When psychometrics and political philosophy meet. Journal of Educational Measurement, *13*, 31-41.
- Feldt, L.S. (1990). The sampling theory for the intraclass reliability coefficient. Applied Measurement in Education, *3*, 361-367.

- Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 105-146). New York: Macmillan.
- Gulliksen, H. (1950). Theory of mental tests. New York: Wiley.
- Haertel, E. H. (1985). Construct validity and criterion-referenced testing. Review of Educational Research, 55, 23-46.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). New York: Macmillan.
- 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.
- Millman, J., & Greene, J. (1989). The specification and development of tests of achievement and ability. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 335-366). New York: Macmillan.
- Petersen, N. S., Kolen, M. J., & Hoover, H. D. (1989). Scaling, norming, and equating. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 221-262). New York: Macmillan.
- Shavelson, R. J., & Webb, N. M. (1991). Generalizability theory: A primer. Newbury Park, CA: Sage.
- Welkowitz, J., Ewen, R.B., Cohen, J. (2000). Introductory statistics for the behavioral sciences (5<sup>th</sup> Ed.). Orlando: Harcourt Brace College.