

ERSH 4600/6600 Applied Educational Assessment
 Spring, 2004: Thursday 5:00 - 7:45
 Aderhold 0409

Instructor: Dr. Deborah Bandalos

Office/phone/email: 325V Aderhold; 542-4540; dbandalo@coe.uga.edu

Office hours: Tuesdays 12:30 - 1:45; Thursdays 3:30-4:45 and by appointment

Text: *Measurement and Evaluation in Psychology and Education*, 6th edition, Robert M. Thorndike.

Supplementary texts: *Measurement and Assessment in Teaching*, 8th edition, Robert L. Linn.

Knowing what Students Know: The Science and Design of Educational Assessment, National Research Council.

Date	Topic	Reading*
1/8	Course Introduction	T: chapter 1
1/15	Introduction to the measurement process; Types of scores	T: chapter 3; Bond, 1996
1/22	Measuring learning outcomes; Objective test items	T: pp. 443-476; NRC, chapter 2
1/29	Criticisms of traditional cognitive testing	Wiggins, 1989; Fredericksen, 1984;
	Cognitive psychology and assessment	NRC: chapter 3
	Measuring complex achievement	L: chapter 9; NRC: chapters 5 & 6
2/5	Essay questions; Performance-based assessments/Portfolios	L: chapter 10; T: chapter 10
2/12	Item analysis: cognitive; review for exam	T: pp. 476-484
2/19	Exam 1	
2/26	Affective measurement	T: chapter 12
3/4	Item analysis for affective measures	
3/11	SPRING BREAK	
3/18	Reliability	T: chapter 4
3/25	Validity	T: chapter 5
4/1	Changing conceptualizations of validity	Hublely & Zumbo, 1996; Anastasi, 1986; Shephard, 1993
Date	Topic	Reading*

4/8	Legal and equity issues	T: chapter 13; Bersoff, 1981;
	1984 Ethics and social considerations 1981	T: chapter 14; Haney,
	Exam 2 due	
4/15	Aptitude tests	T: chapter 8
4/22	Interest and personality assessment	T: chapter 11
	Standardized achievement tests	T: chapter 9
4/29	Project 2 presentations; Project 3 due	

* I reserve the right to assign other readings at my discretion.

T = Thorndike; NRC: National Research Council; L: Linn

EVALUATION

The final grades for this course will be based on a weighted average of grades on exams and projects. Exam 1 will be composed of questions submitted by students. I will review all questions submitted and select questions on the basis of technical merit as discussed in class. The second exam will be a take-home exam. The three major projects are discussed below. However, additional smaller projects or homework assignments may be assigned at my discretion.

Weights for final grades will be as follows:

Exams	40% (20% each)
Projects	60% (Project 1: 10%; Projects 2 & 3: 25% each)

In order to pass the course, all projects must be completed.

GENERAL COURSE GUIDELINES

Much of class time will be devoted to a discussion of the readings. Therefore it is essential that all students read and come prepared to discuss the readings assigned for that day. Students may be called on randomly to comment on the day's readings.

Because much of the work for this class will be done in a group format, it is important for students to be conscientious in guarding against activities that could be interpreted as cheating behavior. Activities such as working in groups, discussing one's work with other students, and sharing ideas **are encouraged** in this class and are **not** considered to constitute cheating. Cheating is defined as copying work from another student, either past or current, or any misrepresentation of another's work as your own. **This includes copying from class notes, readings, or texts on any assignment or exam.** If you are asked for an evaluation or opinion on an assignment or exam, you should give it **in your own words**. Students found to engage in cheating behavior will receive a grade of zero for the assignment on which the cheating occurred and possibly a failing grade in the course. In addition, the behavior will be reported to the student's advisor and other University authorities.

PROJECTS

The projects in this class have been designed to give you hands-on experiences with writing and evaluation different types of measurement. I hope that this will provide a basis for class discussion as we observe and discuss issues arising from these measurement projects. Detailed grading protocols for each project will be provided to you.

PROJECT 1

Project 1 is a group project in which you will develop an assessment of a complex higher order thinking process in some content area. Examples of such assessments are given in Linn, chapter 9 and in chapters 3, 5, and 6 of the NRC book. To keep this project manageable, your group should carefully delineate the content area and cognitive processes of interest. The assessment should be targeted to the specific content and cognitive areas you have delineated.

PROJECT 2

Project 2 is a group project in which you will actually construct and evaluate a short attitude scale. Each group will decide on an area for measurement and group

members will write items for the scale. These items should reflect good item writing principles as discussed in class. In addition to these items, the questionnaire should include four or five items designed to elicit demographic information, such as age or gender. Upon completion of the questionnaire, each group member will be required to administer it. Each group should plan to administer their scale to at least 50 subjects. The subjects need not be chosen randomly.

Each group will be required to complete a set of analyses using the data that is collected on their group's scale. Different sections of the project will be turned in throughout the semester, as the material is covered in class. Groups may redo any portion for which they do not receive all the points possible. Each group is also required to prepare a poster showing their results. These will be shared with the class on the final day of class.

PROJECT 3

Project 3 has been designed to allow you to apply what you have learned in class about test properties by evaluating a standardized test of your choice. At the beginning of the semester, each student will choose a type of test in which they are interested (e.g., intelligence tests, achievement tests, personality tests, etc). The available instruments in this area should then be researched and each student should chose one instrument to evaluate. Project 3 will consist of an assessment of the instrument in terms of reliability, validity, etc. You must therefore chose an instrument on which such information is available.

In order to assess the instrument you will need to obtain a copy of the test manual, and ideally of the test itself. You may be able to obtain these from professors in your department or schools or agencies with whom you are working. They are also available from testing companies, but this will require that you plan ahead as it will take time to order an receive them. Other sources of information include:

the Buros Mental Measurements Yearbooks (these are now available online)

specialty area Tests and Reviews

your textbook and other textbooks available from me.

A written report, no longer than ten typewritten, double-spaced pages will be due the last day of class.

Readings

Anastasi, A. (1986). Evolving concepts of test validation. *Annual Review of Psychology*, 37, 1-15.

Bersoff, D. (1981). Testing and the law. *American Psychologist*, 36(10), 1047-

1056.

Bersoff, D. (1984). Social and legal influences on test development and usage. In Barbara S. Plake (Ed.), *Social and Technical Issues in Testing: Implications for Test Construction and Use*, pp. 87-110. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bond, L.A. (1996). Norm- and criterion-referenced testing. *Practical Assessment, Research, and Evaluation*, 5(2).

Fredericksen, N. (1984). The real test bias: Influences of testing on teaching and learning. *American Psychologist*, 39(3), 193-202.

Hambleton, R. & Rodgers, J. (1995). Item bias review. *Practical Assessment, Research, and Evaluation*, 4(6).

Haney, W. (1981). Validity, vaudeville, and values: A short history of social concerns over standardized testing. *American Psychologist*, 36(10), 1021-1034.

Huble, A.M. & Zumbo, B.D. (1996). A dialectic on validity: Where we have been and where we are going. *The Journal of General Psychology*, 123, 207-215.

National Research Council. (2001). The nature of assessment and reasoning from evidence. In *Knowing what Students Know: The Science and Design of Educational Assessment*, pp. 37-54. Washington, D.C.: National Academy Press.

National Research Council. (2001). Advances in the sciences of thinking and learning. In *Knowing what Students Know: The Science and Design of Educational Assessment*, pp. 59-110. Washington, D.C.: National Academy Press.

National Research Council. (2001). Implications of the new foundations for assessment design. In *Knowing what Students Know: The Science and Design of Educational Assessment*, pp. 177-220. Washington, D.C.: National Academy Press.

National Research Council. (2001). Assessment in practice. In *Knowing what Students Know: The Science and Design of Educational Assessment*, pp. 221-260. Washington, D.C.: National Academy Press.

Rudner, L. (1994). Questions to ask when evaluating tests. *Practical Assessment, Research, and Evaluation*, 4(2).

Shepard, L.A. (1993). Evaluating test validity. *Review of Research in Education*,

19, 405-450.

Wiggins, G. (May, 1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 703-713.